This article was downloaded by: [131.94.16.10] On: 29 September 2015, At: 21:30
Publisher: Institute for Operations Research and the Management Sciences (INFORMS)
INFORMS is located in Maryland, USA

Mathematics of Operations Research

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

A Centered Projective Algorithm for Linear Programming

Michael J. Todd, Yinyu Ye,

To cite this article:
Michael J. Todd, Yinyu Ye, (1990) A Centered Projective Algorithm for Linear Programming. Mathematics of Operations Research 15(3):508-529. http://dx.doi.org/10.1287/moor.15.3.508

Full terms and conditions of use: http://pubsonline.informs.org/page/terms-and-conditions

This article may be used only for the purposes of research, teaching, and/or private study. Commercial use or systematic downloading (by robots or other automatic processes) is prohibited without explicit Publisher approval, unless otherwise noted. For more information, contact permissions@informs.org.

The Publisher does not warrant or guarantee the article's accuracy, completeness, merchantability, fitness for a particular purpose, or non-infringement. Descriptions of, or references to, products or publications, or inclusion of an advertisement in this article, neither constitutes nor implies a guarantee, endorsement, or support of claims made of that product, publication, or service.

© 1990 INFORMS

Please scroll down for article—it is on subsequent pages

INFORMS is the largest professional society in the world for professionals in the fields of operations research, management science, and analytics.
For more information on INFORMS, its publications, membership, or meetings visit http://www.informs.org

MATHEMATICS OF OPERATIONS RESEARCH
Vol. 15, No 3, August 1990
Printed in U S A

A CENTERED PROJECTIVE ALGORITHM FOR LINEAR PROGRAMMING*†

MICHAEL J. TODD AND YINYU YE‡

Cornell University

We describe a projective algorithm for linear programming that shares features with Karmarkar's projective algorithm and its variants and with the path-following methods of Gonzaga, Kojima-Mizuno-Yoshise, Monteiro-Adler, Renegar, Vaidya and Ye. It operates in a primal-dual setting, stays close to the central trajectories, and converges in $O(\sqrt{n}L)$ iterations like the latter methods. (Here n is the number of variables and L the input size of the problem.) However, it is motivated by seeking reductions in a suitable potential function as in projective algorithms, and the approximate centering is an automatic byproduct of our choice of potential function.

1. Introduction. This paper is concerned with interior algorithms for the linear programming problem

$$(P_1)\qquad \begin{aligned} &\min g^{\mathrm{T}}x \\ &Ax = b, \\ &x \geqslant 0, \end{aligned}$$

where A is a real $m \times n$ matrix, b is in R^m, and g and x are in R^n. Apart from methods related to the ellipsoid algorithm, we can distinguish two classes of such methods:

(i) potential or objective function-driven discrete step algorithms: Karmarkar's method [14] and its variants including the affine-scaling variant; and

(ii) path-following algorithms, due to Gonzaga [12], Kojima, Mizuno and Yoshise [15], [16], Monteiro and Adler [21], [22], Renegar [23], Vaidya [28] and Ye [30]. We describe here a method that shares features with both approaches.

Assume that the data A, b and g of (P_1) are integer, and the input size is L. Then Karmarkar's algorithm can obtain within $O(nL)$ iterations a sufficiently accurate solution that an exact solution can easily be deduced. The work required in each iteration is $O(n^3)$ arithmetic operations in the basic version, or an average of $O(n^{2.5})$ operations in the modified algorithm, giving a complexity of $O(n^{3.5}L)$. Each iteration seeks a constant reduction in Karmarkar's potential function; after making a projective transformation, a step is taken in a direction that is the negative of the projected gradient of the potential function in the transformed space. This constant reduction implies the bound of $O(nL)$ for the number of iterations. In practice, far fewer iterations are required; it seems that $O(L)$, or perhaps $O((\ln n)L)$, suffice when a

*Received February 16, 1988; revised March 20, 1989.
AMS 1980 subject classification. Primary: 90C05.
IAOR 1973 subject classification. Main: Programming: linear.
OR/MS Index 1978 subject classification. Primary: 643 Programming/Linear/Algorithms.
Key words. Linear programming, projective algorithms, path following.
†Support in part by the Cowles Foundation for Research in Economics, Yale University, NSF Grant ECS-8602534, ONR Contract N00014-87-K-0212, and the U.S. Army Research Office through the Mathematical Science Institute of Cornell University.
‡Current address: Department of Management Sciences, The University of Iowa.

0364-765X/90/1503/0508/$01.25
Copyright © 1990, The Institute of Management Sciences/Operations Research Society of America

Downloaded from informs.org by [131.94.16.10] on 29 September 2015, at 21:31 . For personal use only, all rights reserved.

Copyright © 2001 All Rights Reserved

RIGHTSLINK

reasonable line search is performed. On the other hand, it is known that a reduction of the potential function greater than some fixed constant cannot always be achieved; see Anstreicher [2] and McDiarmid [17].

Considering the infinitesimal step version of Karmarkar's algorithm and its variants, one is led to study of paths or trajectories in the interior of the feasible region. These paths have been investigated by Bayer and Lagarias [5], Megiddo [18] and Megiddo and Shub [19]. They are closely related to the paths associated with the classical barrier function method of Frisch [11] and Fiacco and McCormick [10], and also to the notion of "analytic center" of Sonnevend [24]. Renegar [23] gave an algorithm using Newton's method to trace the path that required $O(\sqrt{n}\,L)$ iterations, each needing $O(n^3)$ arithmetic operations. Using Karmarkar's trick of solving approximate systems of equations and making rank 1 updates, Vaidya [28] showed that an average of $O(n^{2.5})$ arithmetical operations per iteration sufficed, giving an overall complexity of $O(n^3L)$. Gonzaga [12] independently and simultaneously obtained the same result. All these methods operated in the primal space alone.

Kojima, Mizuno and Yoshise [15] described a primal-dual interior algorithm that followed central trajectories in both the primal and dual feasible regions. This method used $O(nL)$ iterations. Monteiro and Adler [21] gave a primal-dual method that required only $O(\sqrt{n}\,L)$ iterations, and had a complexity of $O(n^3L)$. A similar complexity was established by Kojima et al. [16] for a path-following method for certain linear complementarity problems, including those arising from linear and convex quadratic programming. Monteiro and Adler also extended their method to one for convex quadratic programming with complexity $O(n^3L)$ [22]. All these methods work in a primal-dual setting, although there is no additional computation compared to a primal-only method. Finally, Ye [30] has given a primal-only convex quadratic programming algorithm that requires only $O(\sqrt{n}\,L)$ iterations.

We aim to shed light on these two classes of algorithms by introducing and analyzing a centered projective algorithm for linear programming. Like Karmarkar's method, it is motivated by seeking a reduction in a suitable potential function, and the search direction is the negative of the projected gradient of such a function in a transformed space. Like the path-following methods, it requires only $O(\sqrt{n}\,L)$ iterations, operates in a primal-dual setting, and stays close (automatically) to the central trajectories as long as a fixed step size is used. We analyze our method with exact projections, so that it needs $O(n^3)$ arithmetic operations at each iteration. However, we shall see that the direction generated coincides with those of Monteiro-Adler and Kojima et al., so their analysis shows that inexact projections could be employed to give an overall complexity of $O(n^3L)$. We note that Ye and Todd [32] have shown that the path-following algorithms maintain a reasonable decrease in a certain potential function; this function differs slightly from the one chosen in this paper.

One question we particularly wish to illuminate is the following. The projective methods define a search direction, and the analysis then shows that a constant reduction in the potential function can be achieved; as we have indicated, typical behavior in practice is much more encouraging. On the other hand, the path-following methods try for an *a priori* determined reduction of $(1-\gamma/\sqrt{n})$ in the duality gap for some constant γ and that is what they obtain. Can one hope to do better, and where does the $\sqrt{n}$ term come from? Some suggestions have been given in earlier papers; we hope our approach yields addition insight.

The paper is organized as follows. In §2 we describe Karmarkar's algorithm briefly and indicate why a constant reduction in the potential function is obtained. We also outline the step determination procedure of a path-following method. §3 reformulates the primal and dual problems in a symmetric way and introduces a combined primal-dual problem.

Downloaded from informs.org by [131.94.16.10] on 29 September 2015, at 21:31 . For personal use only, all rights reserved.

Copyright © 2001 All Rights Reserved

RIGHTSLINK

In §4 we state the potential functions with which we will be concerned and discuss scalings of the primal and dual variables. §5 computes the necessary projection matrix and hence the search direction used. We also relate this search direction to the step employed in path-following methods. As is perhaps not too surprising in view of earlier analyses by Gonzaga [13] and Mitchell and Todd [20], our direction is a linear combination of the same two directions that arise in path-following algorithms. However, we stress that the motivation behind it is completely different.

§6 describes the algorithm which is analyzed in §7. Finally §8 contains further discussion.

Since this paper was first written, Ye [31] has obtained a potential function-driven method requiring only $O(\sqrt{n}\,L)$ iterations without the necessity of staying close to the central trajectories. This method is based on the primal-dual potential function studied in this paper.

2. An outline of projective and path-following methods. Our problem is

$$(P_1)\qquad \begin{aligned} &\min g^{\mathrm{T}}x \\ &Ax = b \\ &x \geqslant 0, \end{aligned}$$

where A is $m \times n$, b an m-vector, and g and x lie in R^n. We assume throughout:

(A1) (P_1) has a strictly positive feasible solution;

(A2) The set of optimal solutions of (P_1) is nonempty and bounded; and

(A3) The matrix A has full row rank m.

To describe (the standard-form variant of) Karmarkar's projective algorithm, we make a further simplifying assumption:

(A4) The optimal value of (P_1) is zero and its feasible region contains the strictly positive vector $x^0 = e$, where $e \in R^n$ is a vector of ones, with objective value $g^{\mathrm{T}}e > 0$.

Then it can be shown that (P_1) is equivalent to the homogeneous problem

$$(HP_1)\qquad \begin{aligned} &\min g^{\mathrm{T}}x \\ &[A, -b]\begin{pmatrix} x \\ \xi \end{pmatrix} = 0, \\ &x \geqslant 0, \quad \xi \geqslant 0, \qquad (x, \xi) \neq 0. \end{aligned}$$

For any problem (P), let $F(P)$ denote the feasible region of (P) and $F_+(P)$ the set of strictly positive feasible solutions. We can evaluate any point $(x, \xi) \in F_+(HP_1)$ by Karmarkar's potential function

$$(2.1)\qquad \begin{aligned} \phi_1(x, \xi) &= (n+1)\ln g^{\mathrm{T}}x - \sum_j \ln x_j - \ln \xi \\ &= \sum_j \ln\left(\frac{g^{\mathrm{T}}x}{x_j}\right) + \ln\left(\frac{g^{\mathrm{T}}x}{\xi}\right). \end{aligned}$$

Clearly, ϕ_1 is homogeneous of degree 0, and hence can be viewed as defined on positive rays, i.e., on certain points in projective space.

Downloaded from informs.org by [131.94.16.10] on 29 September 2015, at 21:31 . For personal use only, all rights reserved.

Copyright © 2001 All Rights Reserved

RIGHTSLINK

Karmarkar's algorithm generates a sequence $\{(x^k, \xi^k)\}$ in $F_+(HP_1)$ starting with $(x^0, \xi^0) = (e, 1)$, with

$$\phi_1(x^{k+1}, \xi^{k+1}) \leqslant \phi_1(x^k, \xi^k) - \epsilon \tag{2.2}$$

for some fixed positive ϵ and all k. Without loss of generality, we can take $\xi^k \equiv 1$, and then $\{x^k\}$ is a sequence in $F_+(P_1)$. We can then deduce from (2.2) that

$$g^{\mathrm{T}}x^k \leqslant \left(\frac{e^{\mathrm{T}}x^k}{n}\right)\exp\left(-\frac{k\epsilon}{n+1}\right)g^{\mathrm{T}}x^0 \tag{2.3}$$

where e denotes the vector of ones in R^n. If $F(P_1)$ is bounded, we easily see that $g^{\mathrm{T}}x^k$ converges linearly to the optimal value of zero; using (A2) it can be shown that this holds even if $F(P_1)$ is unbounded.

If the data A, b and g are integer and the total input size is L, (2.3) can be used to show that $O(nL)$ iterations suffice to give an approximate solution from which an exact solution can be obtained by solving a system of equations. In general, $O(nq)$ iterations reduce the objective function by a factor of 2^q.

We now describe how the reduction in potential given by (2.2) is achieved. Let $\hat{x}$ in $F_+(P_1)$ be given, and let $\hat{X} := \operatorname{diag}(\hat{x})$ be the diagonal matrix whose diagonal entries are the components of $\hat{x}$. Let $\bar{g} := \hat{X}g$ and $\bar{A} := A\hat{X}$, and consider the rescaled problem

$$(\overline{HP}_1) \qquad \begin{aligned} &\min \bar{g}^{\mathrm{T}}x \\ &[\bar{A}, -b]\begin{pmatrix}\bar{x}\\ \bar{\xi}\end{pmatrix} = 0, \\ &\bar{x} \geqslant 0, \quad \bar{\xi} \geqslant 0, \qquad (\bar{x}, \bar{\xi}) \neq 0, \end{aligned}$$

in terms of the rescaled variables $\bar{x} = \hat{X}^{-1}x$, $\bar{\xi} = \xi$. It is easy to see that, if (x, ξ) and $(\bar{x}, \bar{\xi})$ correspond as above, then

$$\begin{aligned} \phi_1(x, \xi) - \ln\det \hat{X} &= \bar{\phi}_1(\bar{x}, \bar{\xi}) \\ &:= (n+1)\ln \bar{g}^{\mathrm{T}}\bar{x} - \sum_j \ln \bar{x}_j - \ln \bar{\xi}, \end{aligned} \tag{2.4}$$

and that $(\hat{x}, 1)$ corresponds to the transformed vector $(e, 1)$. From (2.4) it is sufficient to obtain a constant reduction in $\bar{\phi}_1$ from the solution $(e, 1)$. We do this by making a step in the direction given by the negative of the projected gradient of $\bar{\phi}_1$. We find

$$\nabla\bar{\phi}_1 = \nabla\bar{\phi}_1(e, 1) = \begin{pmatrix} \dfrac{n+1}{\bar{g}^{\mathrm{T}}e}\bar{g} - e \\ -1 \end{pmatrix}. \tag{2.5}$$

If P_M denotes projection into the null space of a matrix M, then our direction is

$$\begin{aligned} \bar{d} &:= -P_{[\bar{A}, -b]}\nabla\bar{\phi}_1 \\ &= -P_{[\bar{A}, -b]}P_{[e^{\mathrm{T}}, 1]}\begin{pmatrix} \dfrac{n+1}{\bar{g}^{\mathrm{T}}e}\bar{g} \\ 0 \end{pmatrix} \\ &= -\frac{n+1}{\bar{g}^{\mathrm{T}}e}P_{\begin{bmatrix}\bar{A} & -b \\ e^{\mathrm{T}} & 1\end{bmatrix}}\begin{pmatrix}\bar{g} \\ 0\end{pmatrix}. \end{aligned} \tag{2.6}$$

Copyright © 2001 All Rights Reserved

Downloaded from informs.org by [131.94.16.10] on 29 September 2015, at 21:31 . For personal use only, all rights reserved.

If we move from $(e, 1)$ to $(e, 1) + \beta\bar{d}/\|\bar{d}\|$, we find

$$\bar{\phi}_1((e,1) + \beta\bar{d}/\|\bar{d}\|) = \bar{\phi}_1(e,1) + \beta\bar{d}^{\mathrm{T}}\nabla\bar{\phi}_1/\|\bar{d}\| + O(\beta^2) \tag{2.7}$$
$$= \bar{\phi}_1(e,1) - \beta\|\bar{d}\| + O(\beta^2).$$

It is clear that a constant stepsize β can be chosen without violating positivity. Hence, since the higher order terms can be controlled, the reduction in $\bar{\phi}_1$ will be of the same order as the Euclidean norm of $\bar{d}$. In the proof of their Lemma 3.1, Todd and Burrell [27] show that $\|\bar{d}\| \geqslant 1$. This yields a constant decrease in $\bar{\phi}_1$ per iteration, and hence convergence in $O(nL)$ iterations. Greater reduction in $\bar{\phi}_1$, and hence faster convergence, would follow if $\|\bar{d}\|$ could be shown to increase with n. We will return to this question in §4.

After making a step in projective space to get a new $(\bar{x}, \bar{\xi})$, we normalize to $(\bar{x}/\bar{\xi}, 1)$ (note that $\bar{\phi}_1$ is unchanged), and then rescale to get our new feasible solution $\hat{X}\bar{x}/\bar{\xi}$ to (P_1). Karmarkar's algorithm iterates this procedure.

An alternative method, the so-called affine variant (first proposed by Dikin [6] in 1967; see also [7], Barnes [3] and Vanderbei, Meketon and Freedman [29]), works directly in the original affine space. The rescaled variable $\bar{x} = \hat{X}^{-1}x$ is moved from e in the direction

$$\bar{d}_x = -P_{\bar{A}}\bar{g}. \tag{2.8}$$

Once again, we can use a constant stepsize β in the update $\bar{x} \leftarrow e + \beta\bar{d}_x/\|\bar{d}_x\|$, and this will yield a decrease proportional to $\|\bar{d}_x\|$ in the objective function value. However, no lower bounds on $\|\bar{d}_x\|$ have been established and polynomial convergence has not been shown for this variant (and is believed unlikely; see Megiddo and Shub [19]). We will see (equation (3.8)) that the situation is ameliorated in a primal-dual setting.

Before we turn to path-following methods, we remark on a drawback of the potential function ϕ_1. Assume that x is converging to the unique optimal solution x^* to (P_1), and assume that this is a nondegenerate basic feasible solution. Let $\mathscr{B}$ $(\overline{\mathscr{B}})$ index the basic (nonbasic) components of x^*. Let $\hat{g}$ denote the reduced costs corresponding to the optimal basis, so that $\hat{g}_j$ is positive for $j \notin \overline{\mathscr{B}}$. Then

$$\phi_1(x,1) = (n+1)\ln g^{\mathrm{T}}x - \sum_j \ln x_j$$
$$= (m+1)\ln g^{\mathrm{T}}x + \sum_{j\in\overline{\mathscr{B}}} \ln\left(\frac{\hat{g}^{\mathrm{T}}x}{x_j}\right) - \sum_{j\in\mathscr{B}} \ln x_j$$

(there will be m basic indices); as $x \to x^*$, each term $\ln(\hat{g}^{\mathrm{T}}x/x_j) \geqslant \ln \hat{g}_j$ $(j \in \overline{\mathscr{B}})$ is bounded below, and each term $\ln x_j$ $(j \in \mathscr{B})$ is converging to $\ln x_j^* > -\infty$. Hence a constant reduction in ϕ_1 implies an asymptotic average reduction by a factor $(1 - \gamma/(m+1))$ for some constant γ in the objective function. However, this result is very dependent on primal and dual nondegeneracy. It is not clear that the factor $n + 1$ multiplying $\ln g^T x$ gives the optimal balance as far as complexity is concerned between reducing the objective function and staying away from the constraint boundaries; and finding the right balance might necessitate knowing the degree of primal and dual degeneracy at the solution. We will see that this disadvantage vanishes for our method; indeed, ρ in (4.2) precisely controls this balance.

We now turn to an outline of path-following methods. We discuss a symmetric primal-dual variant, as in [21]. Associated with problem (P_1) is the barrier function

Downloaded from informs.org by [131.94.16.10] on 29 September 2015, at 21:31 . For personal use only, all rights reserved.

Copyright © 2001 All Rights Reserved

RIGHTSLINK

problem with parameter μ:

$$(BP_1)\qquad \begin{aligned} &\min g^{\mathrm{T}}x - \mu \sum_j \ln x_j, \\ &Ax = b, \\ &x > 0; \end{aligned}$$

under our assumptions, this will have a unique solution $x(\mu)$ for each $\mu > 0$, and $x(\mu)$ converges to an optimal solution of (P_1) as $\mu \to 0+$. The optimality conditions for (BP_1), with Lagrange multipliers y, can be written as

$$\begin{aligned} g - \mu X^{-1}e - A^{\mathrm{T}}y &= 0, \\ Ax &= b, \\ x &> 0, \end{aligned}$$

where X denotes diag(x), and S below similarly denotes diag(s). Letting $s := \mu X^{-1}e$, we can write these conditions as

$$XSe - \mu e = 0, \tag{2.9a}$$

$$Ax = b, \tag{2.9b}$$

$$A^{\mathrm{T}}y + s = g, \tag{2.9c}$$

$$x > 0, \qquad s > 0. \tag{2.9d}$$

Thus y, with associate slack vector s, is an interior solution to the dual problem

$$(D_1)\qquad \begin{aligned} &\max b^{\mathrm{T}}y, \\ &A^{\mathrm{T}}y + s = g, \\ &s \geqslant 0, \end{aligned}$$

and moreover the duality gap is $g^{\mathrm{T}}x - b^{\mathrm{T}}y = x^{\mathrm{T}}s = e^{\mathrm{T}}XSe = n\mu$. The equations (2.9) are also the optimality conditions for the barrier function problem

$$(BD_1)\qquad \begin{aligned} &\max b^{\mathrm{T}}y + \mu \sum \ln s_j, \\ &A^{\mathrm{T}}y + s = g, \\ &x > 0, \end{aligned}$$

with unique optimal solution $(y(\mu), s(\mu))$. Hence, solving (2.9) for a sequence of parameters μ approaching zero yields points on the curve $\{x(\mu)\}$ in the primal space and on the curve $\{(y(\mu), s(\mu))\}$ in the dual space. These are the (central) pathways of Megiddo [18] and the central trajectories of Bayer and Lagarias [5]. We call (x, y, s) or (x, s) *centered* if (2.9) holds for some μ.

Equation (2.9a) is nonlinear, so we must be content with approximate solutions. Assume we have a point $(\hat{x}, \hat{y}, \hat{s})$ satisfying (2.9b)–(2.9d) with

$$\|\hat{X}\hat{S}e - \hat{\zeta}e\| \leqslant \alpha\hat{\zeta} \tag{2.10}$$

Copyright © 2001 All Rights Reserved

Downloaded from informs.org by [131.94.16.10] on 29 September 2015, at 21:31 . For personal use only, all rights reserved.

where $\hat{\zeta} := e^{\mathrm{T}}\hat{X}\hat{S}e/n$ and suitable $0 < \alpha < 1$; such a point we will call *approximately centered*. We then seek an approximate solution to (2.9) for some $\mu < \hat{\zeta}$ by taking a Newton step. The new point will be $(\hat{x} + d_x, \hat{y} + d_y, \hat{s} + d_s)$ where the direction satisfies the equations

$$\begin{aligned} \hat{S}d_x + \hat{X}d_s &= -(\hat{X}\hat{S}e - \mu e), \\ Ad_x &= 0, \\ A^{\mathrm{T}}d_y + d_s &= 0. \end{aligned} \tag{2.11}$$

It can be shown (see, e.g., Monteiro and Adler [21]) that the result is also approximately centered, with duality gap approximately $n\mu$, if we choose $\mu = (1 - \gamma/\sqrt{n})\hat{\zeta}$ for some constant γ.

Most of the cited papers give little or no motivation or justification for the choice of μ above. However, Gonzaga [13] in his closely-related primal-only method shows in §3.3 that such a μ for an appropriate γ gives a constant bound less than 1 on $\|\hat{X}^{-1}d_x\|$, thus ensuring that the new point is feasible. Kojima, Mizuno and Yoshise [15] set $\mu = \sigma\hat{\zeta}$ for some constant $\sigma < 1$ and find that they can only take a step size of order $1/n$, resulting in $O(nL)$ iterations. They also show (Theorem 2.2 and the following remark) that, if one is exactly on the path $\{(x(\mu), y(\mu), s(\mu))\}$ and takes a step in the tangent direction for the maximum feasible distance, then the duality gap is reduced by at least the factor $(1 - 1/\sqrt{n})$.

3. Symmetric primal and dual problems. In this section we reformulate (P_1) and (D_1) into symmetric forms, and define a symmetric combined primal-dual problem. Recall that we have

$$\begin{aligned} \min\ & g^{\mathrm{T}}x, \\ & Ax = b, \\ & x \geqslant 0, \end{aligned} \tag{P_1}$$

satisfying (A1)–(A3). Since we assume (P_1) feasible, b is in the range of A, so that we can write

$$b = Ah \tag{3.1}$$

for suitable h. Then the dual can be written

$$\begin{aligned} \max\ & h^{\mathrm{T}}A^{\mathrm{T}}y, \\ & A^{\mathrm{T}}y + s = g, \\ & s \geqslant 0. \end{aligned} \tag{D_1}$$

We prefer to write this in terms of s alone. Thus let B be a matrix whose rows span the null space of A; by choosing a basis for this space we can assume that B is $p \times n$ with full row rank p, with $m + p = n$. Then

$$A^{\mathrm{T}}y + s = g \quad \text{for some } y \quad \text{iff } Bs = Bg$$

and, moreover, the objective function of (D_1) can be written in terms of s alone since

Downloaded from informs.org by [131.94.16.10] on 29 September 2015, at 21:31 . For personal use only, all rights reserved.

Copyright © 2001 All Rights Reserved

RIGHTSLINK

$h^{\mathrm{T}}A^{\mathrm{T}}y = g^{\mathrm{T}}h - h^{\mathrm{T}}s$. Thus we can write (P_1) and (D_1) as

$$(P)\qquad \begin{aligned} &\min g^{\mathrm{T}}x,\\ &Ax = Ah,\\ &x \geqslant 0; \end{aligned}$$

$$(D)\qquad \begin{aligned} &\min h^{\mathrm{T}}s,\\ &Bs = Bg,\\ &s \geqslant 0, \end{aligned}$$

where the rows of A and B span complementary orthogonal subspaces of R^n, which we denote as

$$A \perp B. \tag{3.2}$$

The duality relation between (P) and (D) is easily deduced from the standard result: feasible solutions x and s have objective values satisfying

$$g^{\mathrm{T}}x + h^{\mathrm{T}}s \geqslant g^{\mathrm{T}}h \tag{3.3}$$

and are optimal if and only if equality holds. Indeed, $x - h$ and $s - g$ lie in the null spaces of A and B respectively, and are therefore orthogonal, so

$$g^{\mathrm{T}}x + h^{\mathrm{T}}s - g^{\mathrm{T}}h = x^{\mathrm{T}}s, \tag{3.4}$$

which is nonnegative for feasible x, s.

The standard symmetric inequality-form linear programming problems are naturally included in our format above. If they are written as max $\tilde{c}^{\mathrm{T}}\tilde{x}$, $\tilde{A}\tilde{x} \leqslant \tilde{b}$, $\tilde{x} \geqslant 0$ and min $\tilde{b}^{\mathrm{T}}\tilde{y}$, $\tilde{A}^{\mathrm{T}}\tilde{y} \geqslant \tilde{c}$, $\tilde{y} \geqslant 0$, then we set

$$\begin{aligned} &A = (\tilde{A}, I), \quad B = (I, -\tilde{A}^{\mathrm{T}}),\\ &g = \begin{pmatrix} -\tilde{c} \\ 0 \end{pmatrix} \quad \text{and} \quad h = \begin{pmatrix} 0 \\ \tilde{b} \end{pmatrix} \end{aligned} \tag{3.5}$$

to get corresponding instances (P) and (D).

We can combine (P) and (D) to get the primal-dual problem

$$(PD)\qquad \begin{aligned} &\min g^{\mathrm{T}}x + h^{\mathrm{T}}s - g^{\mathrm{T}}h,\\ &Ax \qquad - Ah = 0,\\ &\qquad Bs - Bg = 0,\\ &x \geqslant 0, \quad s \geqslant 0, \end{aligned}$$

which has optimal value 0 when feasible. As stated, this problem is separable, but we may replace the objective function by $x^{\mathrm{T}}s$ using (3.4). In this form, it is clear that (PD) is unaffected if we replace g by $\tilde{g}$ and h by $\tilde{h}$ where

$$A\tilde{h} = Ah \quad \text{and} \quad B\tilde{g} = Bg. \tag{3.6}$$

Indeed, we have the useful

Downloaded from informs.org by [131.94.16.10] on 29 September 2015, at 21:31 . For personal use only, all rights reserved.

Copyright © 2001 All Rights Reserved

RIGHTSLINK

Downloaded from informs.org by [131.94.16.10] on 29 September 2015, at 21:31 . For personal use only, all rights reserved.

LEMMA 1. *Let $(\tilde{P})$ and $(\tilde{D})$ denote (P) and (D) defined using $\tilde{g}$ and $\tilde{h}$ in place of g and h, where (3.6) holds. Then $(\tilde{P})$ and $(\tilde{D})$ are equivalent to (P) and (D) in the sense that their feasible regions are unchanged and their objective functions the same up to additive constants.*

PROOF. It is clear that the feasible regions are unchanged. Now (3.6) implies that $\tilde{g} = g + A^{\mathrm{T}}u$ and $\tilde{h} = h + B^{\mathrm{T}}v$ for some u, v. Thus, for any feasible x,

$$\tilde{g}^{\mathrm{T}}x = g^{\mathrm{T}}x + u^{\mathrm{T}}Ax = g^{\mathrm{T}}x + u^{\mathrm{T}}Ah = g^{\mathrm{T}}x + (\tilde{g}^{\mathrm{T}}h - g^{\mathrm{T}}h)$$

differs by a constant from $g^{\mathrm{T}}x$. Similarly, $\tilde{h}^{\mathrm{T}}s$ differs by a constant from $h^{\mathrm{T}}s$ for feasible s.

We will use lemma 1 as follows; given $(\hat{x}, \hat{s}) \in F(PD)$, we can assume without loss of generality that $g = \hat{s}$ and $h = \hat{x}$.

Let us observe immediately one consequence of this. Suppose that $x = s = e$ is feasible in (PD). This solution is also centered, since (2.9a) holds with $\mu = 1$. Then we can take $g = h = e$ also. The direction chosen by the affine algorithm (see (2.8)) is then

$$d = \begin{pmatrix} d_x \\ d_s \end{pmatrix} = -\begin{pmatrix} P_A & 0 \\ 0 & P_B \end{pmatrix}\begin{pmatrix} e \\ e \end{pmatrix} = \begin{pmatrix} -P_A e \\ -P_B e \end{pmatrix}. \tag{3.7}$$

Recall that the difficulty with the affine variant was in bounding from below the norm of $\|\bar{d}_x\|$; even if $-\bar{g}$ were large (e.g., $\bar{g} = e$), its projection $\bar{d}_x$ into the null space of $\bar{A}$ could be small. Here this problem vanishes; since $A \perp B$, $d_x + d_s = -e$ and we can conclude that

$$\|d\| = \sqrt{n}\,. \tag{3.8}$$

Note that centering alone is not enough in this analysis—we need also the primal-dual formulation. In computing d from (3.7), the increased dimension implies no increase in work; given $d_x = -P_A e$, we find $d_s = -P_B e = -e - d_x$, since

$$P_A + P_B = I. \tag{3.9}$$

The importance of centering in the affine variant has been demonstrated also in Barnes, Chopra, and Jensen [4].

Since (PD) has optimal value 0, it is natural to consider its homogenization

$$(HPD) \qquad \begin{aligned} &\min g^{\mathrm{T}}x + h^{\mathrm{T}}s - g^{\mathrm{T}}h\tau, \\ &Ax \qquad\quad - Ah\tau = 0, \\ &\qquad\quad Bs - Bg\tau = 0 \\ &x \geqslant 0, \quad s \geqslant 0, \quad \tau \geqslant 0, \qquad (x, s, \tau) \neq 0. \end{aligned}$$

As for (HP_1), we can view the feasible solutions of (HPD) as rays, i.e., as points in projective space. The relation equivalent to (3.4) for $(x, s, \tau) \in F(HPD)$ with $\tau > 0$ is

$$g^{\mathrm{T}}x + h^{\mathrm{T}}s - g^{\mathrm{T}}h\tau = x^{\mathrm{T}}s/\tau. \tag{3.10}$$

Copyright © 2001 All Rights Reserved

To conclude this section, we observe that our assumptions (A1) and (A2) easily imply (B1) and (B2) below, and (B3) is, like (A3), without loss of generality:

(B1) (D) has a strictly positive feasible solution;

(B2) The set of optimal solutions of (D) is nonempty and bounded; and

(B3) The matrix B has full row rank p.

Hence (PD) has a strictly positive feasible solution, i.e. $F_+(PD) \neq \varnothing$, and also $F_+(HPD) \neq \varnothing$.

4. Potential functions and scaling. We associate with the homogeneous primal-dual problem the potential function

$$\phi(x, s, \tau) = \phi_\rho(x, s, \tau; g, h) \tag{4.1}$$

$$:= (n + \rho)\ln(g^{\mathrm{T}}x + h^{\mathrm{T}}s - g^{\mathrm{T}}h\tau) - \sum \ln x_j - \sum \ln s_j - (\rho - n)\ln \tau$$

on $F_+(HPD)$. Note that this is homogeneous of degree 0, and hence is defined on (positive) rays. Perhaps the most natural choice of ρ is $n + 1$, leading to the potential function introduced by Karmarkar [14]. See also Ye and Todd [32]. However, we will find it profitable to consider other values for ρ, in particular $\rho \sim \sqrt{n}$. While it may seem strange to have a negative coefficient for $\ln \tau$ in (4.1), we note that (3.10) implies

$$\phi(x, s, \tau) = (n + \rho)\ln x^{\mathrm{T}}s - \sum \ln x_j - \sum \ln s_j - 2\rho \ln \tau \tag{4.2}$$

$$= \rho \ln\left(\frac{x^{\mathrm{T}}s}{\tau^2}\right) - \sum \ln \frac{x_j s_j}{x^{\mathrm{T}}s}.$$

This shows that any $\rho > 0$ is reasonable and provides a natural interpretation to the potential function, including the role of ρ. Indeed, $x^{\mathrm{T}}s/\tau^2$ is the objective function of the unscaled solution $(x/\tau, s/\tau)$ in $F_+(PD)$, while $(x_j s_j/x^{\mathrm{T}}s)$ is a positive vector lying on the unit simplex. Hence ρ balances a term that measures the objective function with the barrier function that seeks to keep $(x/\tau, s/\tau)$ centered in $F_+(PD)$.

Our aim is to secure an appropriate decrease in ϕ at each iteration. To this end we will work with scaled problems.

Let Ω be a positive definite diagonal $n \times n$ matrix and δ a positive scalar. Define the scaled data

$$\bar{A} = \delta A\Omega, \quad \bar{B} = \delta B\Omega^{-1}, \tag{4.3}$$

$$\bar{g} = \delta\Omega g, \quad \bar{h} = \delta\Omega^{-1}h.$$

Then, in terms of the rescaled variables

$$\bar{x} = \delta\Omega^{-1}x, \quad \bar{s} = \delta\Omega s, \quad \bar{\tau} = \tau, \tag{4.4}$$

Downloaded from informs.org by [131.94.16.10] on 29 September 2015, at 21:31 . For personal use only, all rights reserved.

Copyright © 2001 All Rights Reserved

RIGHTSLINK

we have the equivalent problem

$$(\overline{HPD})\qquad \begin{aligned} &\min \bar{g}^{\mathrm{T}}\bar{x}+\bar{h}^{\mathrm{T}}\bar{s}-\bar{g}^{\mathrm{T}}\bar{h}\bar{\tau},\\ &\bar{A}\bar{x} \qquad -\bar{A}\bar{h}\bar{\tau}=0,\\ &\qquad \bar{B}\bar{s}\ -\bar{B}\bar{g}\bar{\tau}=0,\\ &\bar{x}\geqslant 0,\quad \bar{s}\geqslant 0,\quad \bar{\tau}\geqslant 0,\qquad (\bar{x},\bar{s},\bar{\tau})\neq 0. \end{aligned}$$

We let $(\overline{PD})$ denote the corresponding inhomogeneous problem, obtained by replacing $\bar{\tau}$ by 1 in $(\overline{HPD})$. It follows from the definition (4.3) that

$$\bar{A}\perp\bar{B}; \tag{4.5}$$

also the objective function of $(\overline{HPD})$ is δ^2 times that of (HPD) for corresponding points. If we set

$$\bar{\phi}(\bar{x},\bar{s},\bar{\tau}) := \phi_\rho(\bar{x},\bar{s},\bar{\tau};\bar{g},\bar{h}) \tag{4.6}$$

Then we find

$$\bar{\phi}(\bar{x},\bar{s},\bar{\tau}) = \phi(x,s,\tau)+\rho\ln\delta^2$$

whenever the arguments correspond as in (4.4). Thus to reduce the original potential function, it suffices to reduce $\bar{\phi}$ sufficiently for some appropriately scaled problem.

We now describe a particularly useful rescaling. Suppose we have $(\hat{x},\hat{s})\in F_+(PD)$, so that $(\hat{x},\hat{s},1)\in F_+(HPD)$ with objective value $\hat{x}^{\mathrm{T}}\hat{s}$. Let

$$\hat{\zeta} := \hat{x}^{\mathrm{T}}\hat{s}/n. \tag{4.7}$$

According to Lemma 1, we may take $g=\hat{s}$, $h=\hat{x}$ without changing the problems.

We also suppose $(\hat{x},\hat{s})$ is approximately centered in $F_+(PD)$, so that

$$\|\hat{X}\hat{S}e-\hat{\zeta}e\|\leqslant\alpha\hat{\zeta} \tag{4.8}$$

for suitable $0<\alpha<1$. Now choose

$$\delta=\hat{\zeta}^{-1/2}\quad\text{and}\quad \Omega=(\hat{X}\hat{S}^{-1})^{1/2}. \tag{4.9}$$

Then, with $g=\hat{s}$, $h=\hat{x}$, we find

$$\bar{g}=\bar{h}=\hat{\zeta}^{-1/2}(\hat{X}\hat{S})^{1/2}e =: \tilde{e}\quad\text{and} \tag{4.10}$$

$$\bar{g}^{\mathrm{T}}\bar{h}=\hat{\zeta}^{-1}e^{\mathrm{T}}\hat{X}\hat{S}e=n. \tag{4.11}$$

Since $\tilde{e}$ is close to e, and our current point $(\hat{x},\hat{s},1)$ has been rescaled to $(\tilde{e},\tilde{e},1)$, we have simultaneously scaled the primal and dual problems reasonably well. At the same time, by employing a symmetric scaling for the primal and dual, we guarantee condition (4.5), which, as we shall see, is crucial.

5. Projections and the search direction. We will be working with the scaled problem $(\overline{HPD})$ and the associated potential function $\bar{\phi}$. The scaling will usually be that at the end of §4, so that $\bar{g}=\bar{h}$, but we first compute the direction in general,

Downloaded from informs.org by [131.94.16.10] on 29 September 2015, at 21:31 . For personal use only, all rights reserved.

Copyright © 2001 All Rights Reserved

assuming only that

$$\bar{g}^{\mathrm{T}}\bar{h} = n \tag{5.1}$$

for notational simplicity; this can always be achieved by a scalar parameter. We also assume that $\bar{g}$ and $\bar{h}$ have been adjusted if necessary so that our current solution is

$$\begin{pmatrix} \bar{h} \\ \bar{g} \\ 1 \end{pmatrix}. \tag{5.2}$$

We first evaluate the gradient of the potential function at the current point. For a vector $u = (u_j) \in R^n$, it is convenient to denote

$$u^{-1} := \left(u_j^{-1}\right) \tag{5.3}$$

and to write $u^{-\mathrm{T}}$ for $(u^{-1})^{\mathrm{T}}$. Then the gradient, using (4.2), is

$$\nabla\bar{\phi} = \nabla\bar{\phi}(\bar{h}, \bar{g}, 1) = \frac{n+\rho}{n}\begin{pmatrix} \bar{g} \\ \bar{h} \\ 0 \end{pmatrix} - \begin{pmatrix} \bar{h}^{-1} \\ \bar{g}^{-1} \\ 2\rho \end{pmatrix}. \tag{5.4}$$

The search direction we will employ is, as in projective methods for (P_1), the projection of $-\nabla\bar{\phi}$ into the tangent space of the feasible region. Hence, with

$$\bar{C} := \begin{bmatrix} \bar{A} & 0 & -\bar{A}\bar{h} \\ 0 & \bar{B} & -\bar{B}\bar{g} \end{bmatrix}, \tag{5.5}$$

we seek

$$\bar{d}' := -P_{\bar{C}}\,\nabla\bar{\phi}. \tag{5.6}$$

An important consequence of the primal-dual setting is that, with the scaling given at the end of §4, $\bar{d}'$ (or its modification $\bar{d}$ below) cannot be too short.

Before computing $\bar{d}'$, we remark that, if the current point has $\bar{g} = \bar{h} = e$, $\tau = 1$, then the homogeneity of $\bar{\phi}$ implies $(e^{\mathrm{T}}, e^{\mathrm{T}}, 1)\nabla\bar{\phi} = 0$. Hence $\bar{d}'$ is also the projection of $-\nabla\bar{\phi}$ into the null space of

$$\begin{bmatrix} \bar{A} & 0 & -\bar{A}\bar{h} \\ 0 & \bar{B} & -\bar{B}\bar{g} \\ e^{\mathrm{T}} & e^{\mathrm{T}} & 1 \end{bmatrix},$$

and this direction is also appropriate if the "simplex constraint" $e^{\mathrm{T}}\bar{x} + e^{\mathrm{T}}\bar{s} + \bar{\tau} = 2n + 1$ is added to $(\overline{HPD})$. A scalar multiple of $\bar{d}'$ is obtained if we project the negative gradient of the objective function into this same null space. However, if we work with $\nabla\bar{\phi}$, the extra appended row is unnecessary. Omitting the simplex constraint preserves the homogeneity of $(\overline{HPD})$ and has been propounded strongly by de Ghellinck and Vial [8], [9]. In our context, adding a row of ones changes the direction since the current solution is generally not $\bar{x} = \bar{s} = e$, $\tau = 1$.

Copyright © 2001 All Rights Reserved

Downloaded from informs.org by [131.94.16.10] on 29 September 2015, at 21:31 . For personal use only, all rights reserved.

We now need to compute $P_{\bar{C}}$. An important fact which follows from (4.5) is that

$$P_{\bar{A}} + P_{\bar{B}} = I; \tag{5.7}$$

we use this repeatedly. Let

$$T := \begin{pmatrix} \bar{A} & 0 \\ 0 & \bar{B} \end{pmatrix}, \quad u := \begin{pmatrix} -\bar{h} \\ -\bar{g} \end{pmatrix}, \quad \text{so } \bar{C} = [T, Tu].$$

Then

$$P_T = \begin{pmatrix} P_{\bar{A}} & 0 \\ 0 & P_{\bar{B}} \end{pmatrix}, \quad I - P_T = \begin{pmatrix} P_{\bar{B}} & 0 \\ 0 & P_{\bar{A}} \end{pmatrix}, \quad \text{so}$$

$$v := (I - P_T)u = \begin{pmatrix} -P_{\bar{B}}\bar{h} \\ -P_{\bar{A}}\bar{g} \end{pmatrix}$$

Hence

$$\omega := (1 + v^{\mathrm{T}}v)^{-1} = \left(1 + \bar{h}^{\mathrm{T}}P_{\bar{B}}\bar{h} + \bar{g}^{\mathrm{T}}P_{\bar{A}}\bar{g}\right)^{-1}. \tag{5.8}$$

Now from §4.2 of [26] we obtain

LEMMA 2.

$$P_{\bar{C}} = \begin{pmatrix} P_{\bar{A}} & 0 & 0 \\ 0 & P_{\bar{B}} & 0 \\ 0 & 0 & 0 \end{pmatrix} + \omega \begin{pmatrix} P_{\bar{B}}\bar{h} \\ P_{\bar{A}}\bar{g} \\ 1 \end{pmatrix} \begin{pmatrix} P_{\bar{B}}\bar{h} \\ P_{\bar{A}}\bar{g} \\ 1 \end{pmatrix}^{\mathrm{T}} \tag{5.9}$$

where ω is given by (5.8).

To compute $\bar{d}'$ in (5.6), we first calculate

$$\begin{pmatrix} P_{\bar{B}}\bar{h} \\ P_{\bar{A}}\bar{g} \\ 1 \end{pmatrix}^{\mathrm{T}} \begin{pmatrix} \bar{g} \\ \bar{h} \\ 0 \end{pmatrix} = \bar{g}^{\mathrm{T}}P_{\bar{B}}\bar{h} + \bar{g}^{\mathrm{T}}P_{\bar{A}}\bar{h} = \bar{g}^{\mathrm{T}}\bar{h} = n$$

using (5.7), and

$$\begin{pmatrix} P_{\bar{B}}\bar{h} \\ P_{\bar{A}}\bar{g} \\ 1 \end{pmatrix}^{\mathrm{T}} \begin{pmatrix} \bar{h}^{-1} \\ \bar{g}^{-1} \\ 2\rho \end{pmatrix} = \bar{h}^{-\mathrm{T}}P_{\bar{B}}\bar{h} + \bar{g}^{-\mathrm{T}}P_{\bar{A}}\bar{g} + 2\rho$$

$$= n + \rho + \sigma, \quad \text{where}$$

$$\sigma := \rho + \bar{g}^{-\mathrm{T}}P_{\bar{A}}\bar{g} + \bar{h}^{-\mathrm{T}}P_{\bar{B}}\bar{h} - n. \tag{5.10}$$

Downloaded from informs.org by [131.94.16.10] on 29 September 2015, at 21:31 . For personal use only, all rights reserved.

Copyright © 2001 All Rights Reserved

RIGHTSLINK

From (5.4), (5.6), (5.9) and the above, we find

$$\bar{d}' = -P_{\bar{C}}\nabla\bar{\phi} = -\frac{n+\rho}{n}\left[\begin{pmatrix} P_{\bar{A}}\bar{g} \\ P_{\bar{B}}\bar{h} \\ 0 \end{pmatrix} + \omega n \begin{pmatrix} P_{\bar{B}}\bar{h} \\ P_{\bar{A}}\bar{g} \\ 1 \end{pmatrix}\right] + \left[\begin{pmatrix} P_{\bar{A}}\bar{h}^{-1} \\ P_{\bar{B}}\bar{g}^{-1} \\ 0 \end{pmatrix} + \omega(n+\rho+\sigma)\begin{pmatrix} P_{\bar{B}}\bar{h} \\ P_{\bar{A}}\bar{g} \\ 1 \end{pmatrix}\right] = -\frac{n+\rho}{n}\begin{pmatrix} P_{\bar{A}}\bar{g} \\ P_{\bar{B}}\bar{h} \\ 0 \end{pmatrix} + \begin{pmatrix} P_{\bar{A}}\bar{h}^{-1} \\ P_{\bar{B}}\bar{g}^{-1} \\ 0 \end{pmatrix} + \omega\sigma\begin{pmatrix} P_{\bar{B}}\bar{h} \\ P_{\bar{A}}\bar{g} \\ 1 \end{pmatrix}. \tag{5.11}$$

The direction $\bar{d}'$ is appropriate to projective space, the domain of $(\overline{HPD})$. It is convenient to transform this direction back to an appropriate direction for the inhomogeneous problem $(\overline{PD})$. If we move from our current point a step size β' in the direction $\bar{d}' = (\bar{d}'_x, \bar{d}'_s, \bar{d}'_\tau)$, we reach

$$\begin{pmatrix} \bar{h} + \beta'\bar{d}'_x \\ \bar{g} + \beta'\bar{d}'_s \\ 1 + \beta'\bar{d}'_\tau \end{pmatrix} = (1 + \beta'\bar{d}'_\tau)\begin{pmatrix} \bar{h} + \beta\bar{d}_x \\ \bar{g} + \beta\bar{d}_s \\ 1 \end{pmatrix}, \quad \text{where}$$

$$\beta := \beta'/(1 + \beta'\bar{d}'_\tau) \quad \text{and}$$

$$\bar{d}_x = \bar{d}'_x - \bar{d}'_\tau\bar{h}, \quad \bar{d}_s = \bar{d}'_s - d'_\tau\bar{g}.$$

Given the homogeneity of $\bar{\phi}$, it is therefore equivalent to move a step size β in the direction

$$\bar{d} = \begin{pmatrix} \bar{d}_x \\ \bar{d}_s \\ \bar{d}_\tau \end{pmatrix} := -\frac{n+\rho}{n}\begin{pmatrix} P_{\bar{A}}\bar{g} \\ P_{\bar{B}}\bar{h} \\ 0 \end{pmatrix} + \begin{pmatrix} P_{\bar{A}}\bar{h}^{-1} \\ P_{\bar{B}}\bar{g}^{-1} \\ 0 \end{pmatrix} + \omega\sigma\begin{pmatrix} -P_{\bar{A}}\bar{h} \\ -P_{\bar{B}}\bar{g} \\ 0 \end{pmatrix}. \tag{5.12}$$

We now note the simplifications that occur when, as in (4.10), $\bar{g} = \bar{h} = \tilde{e}$. Then, from (5.1), $\tilde{e}^{\mathrm{T}}\tilde{e} = n$ and clearly $\tilde{e}^{-\mathrm{T}}\tilde{e} = n$. Hence from (5.8)

$$\omega = \left(1 + \tilde{e}^{\mathrm{T}}P_{\bar{A}}\tilde{e} + \tilde{e}^{\mathrm{T}}P_{\bar{B}}\tilde{e}\right)^{-1} = (1 + \tilde{e}^{\mathrm{T}}\tilde{e})^{-1} = (n+1)^{-1}$$

and from (5.10)

$$\sigma = \rho + \tilde{e}^{-\mathrm{T}}P_{\bar{A}}\tilde{e} + \tilde{e}^{-\mathrm{T}}P_{\bar{B}}\tilde{e} - n = \rho + \tilde{e}^{-\mathrm{T}}\tilde{e} - n = \rho.$$

Downloaded from informs.org by [131.94.16.10] on 29 September 2015, at 21:31 . For personal use only, all rights reserved.

Copyright © 2001 All Rights Reserved

RIGHTSLINK

Thus

$$\bar{d} = -(1+\psi)\begin{pmatrix} P_{\bar{A}}\tilde{e} \\ P_{\bar{B}}\tilde{e} \\ 0 \end{pmatrix} + \begin{pmatrix} P_{\bar{A}}\tilde{e}^{-1} \\ P_{\bar{B}}\tilde{e}^{-1} \\ 0 \end{pmatrix}, \quad \text{where} \tag{5.13}$$

$$\psi = \frac{2n+1}{n(n+1)}\rho. \tag{5.14}$$

Again, if $\tilde{e}$ is close to e and a suitable value of ψ is chosen, our primal-dual setting assures us that $\bar{d}$ cannot be too small. Further, there is no computational expense for this increased dimensionality; if

$$f := -(1+\psi)\tilde{e} + \tilde{e}^{-1}, \tag{5.15}$$

then

$$\bar{d}_x = P_{\bar{A}} f \quad \text{and} \quad \bar{d}_s = f - \bar{d}_x. \tag{5.16}$$

To conclude this section, we relate our direction $\bar{d}$ in (5.13) to those generated by path-following methods. Consider equations (2.11). If we scale as at the end of §4, we can rewrite these equations in terms of $\bar{d}_x = \delta\Omega^{-1}d_x$, $\bar{d}_s = \delta\Omega d_s$. The last equation of (2.11) is equivalent to $Bd_s = 0$. We thus obtain

$$(\hat{X}\hat{S})^{1/2}(\bar{d}_x + \bar{d}_s) = -\delta(\hat{X}\hat{S}e - \mu e),$$

$$\bar{A}\bar{d}_x = 0,$$

$$\bar{B}\bar{d}_s = 0.$$

The first equation can be written

$$\bar{d}_x + \bar{d}_s = -\tilde{e} + (\mu/\hat{\zeta})\tilde{e}^{-1},$$

and so

$$\bar{d}_x = P_{\bar{A}} f', \quad \bar{d}_s = P_{\bar{B}} f', \quad \text{with} \tag{5.17}$$

$$f' = -\tilde{e} + (\mu/\hat{\zeta})\tilde{e}^{-1}. \tag{5.18}$$

The similarity to (5.15), (5.16) is striking. Indeed, if ψ is chosen so that $1 + \psi = \hat{\zeta}/\mu$, then the two directions coincide. If $\mu = \hat{\zeta}(1 - \gamma/\sqrt{n})$ for some constant γ, this suggests $\psi = \gamma'/\sqrt{n}$ for some constant γ'. Our choice of ψ will turn out to have this form, but it is motivated by obtaining suitable reductions in the potential function $\bar{\phi}$.

6. The algorithm. We now have all the ingredients for our algorithm. We assume we have available a point $(x^0, s^0) \in F_+(PD)$ that is approximately centered, so that

$$\|X^0S^0e - \zeta^0 e\| \leqslant \alpha\zeta^0, \tag{6.1}$$

where $\zeta^0 = e^{\mathrm{T}}X^0S^0e/n$, for suitable $0 < \alpha < 1$. Obtaining such a point has been discussed in [12], [15], [16], [21]–[23], [28]. We generate a sequence $\{(x^k, s^k)\} \subseteq F_+(PD)$ as follows.

Copyright © 2001 All Rights Reserved

Downloaded from informs.org by [131.94.16.10] on 29 September 2015, at 21:31 . For personal use only, all rights reserved.

Given (x^k, s^k), set $\hat{x} = x^k$, $\hat{s} = s^k$ and define the scaled data by (4.3), with δ and Ω given by (4.9). Then (x^k, s^k) is transformed into $(\tilde{e}, \tilde{e})$. We choose $\bar{d}$ by (5.13) and set

$$\begin{pmatrix} \bar{x} \\ \bar{s} \end{pmatrix} := \begin{pmatrix} \tilde{e} \\ \tilde{e} \end{pmatrix} + \beta \begin{pmatrix} \bar{d}_x \\ \bar{d}_s \end{pmatrix} \tag{6.2}$$

for appropriate $\beta > 0$. (In fact, we analyze the case of a fixed β. However, a line search to minimize approximately the potential function could be used as long as it included a safeguard to keep the iterates approximately centered. Ye's recent algorithm [31] does not require such a safeguard.) Then we unscale the point to get

$$\begin{pmatrix} x^{k+1} \\ s^{k+1} \end{pmatrix} = \begin{pmatrix} \delta^{-1}\Omega\bar{x} \\ \delta^{-1}\Omega^{-1}\bar{s} \end{pmatrix}. \tag{6.3}$$

We aim to prove

THEOREM 1. *Let*

$$\rho = \frac{2n+2}{2n+1}\sqrt{n} \quad \textit{and} \quad \psi = \frac{2}{\sqrt{n}}.$$

Let $\alpha = 1/3$ and $\beta = 1/15$. Then if (x^0, s^0) satisfies (6.1), *the algorithm above is well-defined and generates a sequence of points satisfying*

$$\|X^k S^k e - \zeta^k e\| \leqslant \alpha \zeta^k \tag{6.4}$$

where $\zeta^k = e^{\mathrm{T}} X^k S^k e/n$ and

$$\phi(x^{k+1}, s^{k+1}, 1) \leqslant \phi(x^k, s^k, 1) - \eta \tag{6.5}$$

where $\eta = 1/9$.

The inequality (6.5) implies the desired complexity of the algorithm. Indeed, from (4.2)

$$\phi(x, s, 1) = \rho \ln(x^{\mathrm{T}} s) - \sum_j \ln \frac{x_j s_j}{x^{\mathrm{T}} s}$$

and (6.1) shows that the last term is bounded for (x^0, s^0) (and by (6.4) for all (x^k, s^k)). Hence, with $\rho \sim \sqrt{n}$, (6.5) gives

COROLLARY 1. *In $O(\sqrt{n}\,L)$ steps, the algorithm above yields*

$$e^{\mathrm{T}} X^k S^k e \leqslant 2^{-L} e^{\mathrm{T}} X^0 S^0 e. \tag{6.6}$$

In the integer model, a suitable choice for (x^0, s^0) will then show that $O(\sqrt{n}\,L)$ steps suffice to give a solution from which optimal solutions to (P) and (D) can be recovered in $O(n^3)$ arithmetic operations. See [12], [15], [16], [21]–[23], [28]. Here L is the input size of the instance.

In fact, our analysis below will show (6.6) directly, by proving that the objective function decreases by a factor $(1 - \gamma/\sqrt{n})$ for constant γ at each iteration. However, since our algorithm is motivated by decreasing the potential function ϕ, we prefer to use it as a criterion. The approximate centering shown by (6.4) is an automatic byproduct of the method with constant step sizes.

Downloaded from informs.org by [131.94.16.10] on 29 September 2015, at 21:31 . For personal use only, all rights reserved.

Copyright © 2001 All Rights Reserved

Towards proving Theorem 1, we note that conditions (6.4)–(6.5) are invariant under scaling of the problem, so that it is sufficient to show that the move from $(\tilde{e}, \tilde{e})$ to $(\bar{x}, \bar{s})$ preserves (6.4) and yields the potential function reduction (6.5). Removing the overbars for notational simplicity, we only need to establish the following (here $\tilde{e}^2$ denotes the vector with $(\tilde{e}^2)_j = (\tilde{e}_j)^2$):

LEMMA 3. *Let ρ, ψ, α, β and η be as in the theorem. Suppose*

$$\|\tilde{e}^2 - e\| \leqslant \alpha. \tag{6.7}$$

If

$$d = \begin{pmatrix} d_x \\ d_s \\ d_\tau \end{pmatrix} = -(1+\psi)\begin{pmatrix} P_A\tilde{e} \\ P_B\tilde{e} \\ 0 \end{pmatrix} + \begin{pmatrix} P_A\tilde{e}^{-1} \\ P_B\tilde{e}^{-1} \\ 0 \end{pmatrix} \quad \textit{and} \tag{6.8}$$

$$\begin{pmatrix} x \\ s \end{pmatrix} := \begin{pmatrix} \tilde{e} \\ \tilde{e} \end{pmatrix} + \beta\begin{pmatrix} d_x \\ d_s \end{pmatrix},$$

then

$$\|XSe - \zeta e\| \leqslant \alpha\zeta \tag{6.9}$$

with $\zeta = x^{\mathrm{T}}s/n$, and

$$\phi(x, s, 1) \leqslant \phi(\tilde{e}, \tilde{e}, 1) - \eta. \tag{6.10}$$

7. Analysis. This section proves Lemma 3. We carry out the analysis as far as possible using general values for ρ, ψ, α, β and η. This will help us discuss the choice of ρ, ψ made. Only when we complete the proof of Lemma 3 will the specific values be used. However, we assume throughout that β is such that $(x, s) > 0$. Since $(\tilde{e}, \tilde{e})$ is feasible and d_x (d_s) in the null space of A (B), this assures us that (x, s) is feasible also.

We now calculate some key quantities. As usual, we denote by D_x and D_s $\mathrm{diag}(d_x)$ and $\mathrm{diag}(d_s)$ respectively, and $\tilde{E}$ denotes $\mathrm{diag}(\tilde{e})$.

LEMMA 4. *We have*

$$-\nabla\phi^{\mathrm{T}}d = \rho\psi + \tilde{e}^{-\mathrm{T}}\tilde{e}^{-1} - n, \tag{7.1}$$

$$d^{\mathrm{T}}d = n\psi^2 + \tilde{e}^{-\mathrm{T}}\tilde{e}^{-1} - n, \tag{7.2}$$

$$\zeta := x^{\mathrm{T}}s/n = 1 - \beta\psi, \quad \textit{and} \tag{7.3}$$

$$XSe - \zeta e = (1 - \beta(1+\psi))(\tilde{e}^2 - e) + \beta^2 D_x D_s e. \tag{7.4}$$

PROOF. In the present setting, (5.4) simplifies to

$$\nabla\phi = \frac{n+\rho}{n}\begin{pmatrix} \tilde{e} \\ \tilde{e} \\ 0 \end{pmatrix} - \begin{pmatrix} \tilde{e}^{-1} \\ \tilde{e}^{-1} \\ 2\rho \end{pmatrix}.$$

Downloaded from informs.org by [131.94.16.10] on 29 September 2015, at 21:31 . For personal use only, all rights reserved.

Copyright © 2001 All Rights Reserved

Then (7.1) and (7.2) are straightforward consequences of (6.8), using $P_A + P_B = I$ repeatedly.

Next note that $d_x^{\mathrm{T}} d_s = 0$, since d_x and d_s lie in the null spaces of A and B respectively, while

$$d_x + d_s = -(1+\psi)\tilde{e} + \tilde{e}^{-1}. \tag{7.5}$$

Hence

$$\begin{aligned}\zeta &= (\tilde{e} + \beta d_x)^{\mathrm{T}}(\tilde{e} + \beta d_s)/n \\ &= \left(\tilde{e}^{\mathrm{T}}\tilde{e} + \beta\tilde{e}^{\mathrm{T}}(d_x + d_s)\right)/n \\ &= (n + \beta(-n(1+\psi) + n))/n \\ &= 1 - \beta\psi,\end{aligned}$$

giving (7.3), and

$$\begin{aligned}XSe - \zeta e &= \left(\tilde{E} + \beta D_x\right)\left(\tilde{E} + \beta D_s\right)e - \zeta e \\ &= \tilde{e}^2 + \beta\tilde{E}(d_x + d_s) + \beta^2 D_x D_s e - (1 - \beta\psi)e \\ &= \tilde{e}^2 + \beta\left(-(1+\psi)\tilde{e}^2 + e\right) + \beta^2 D_x D_s e - (1 - \beta\psi)e \\ &= (1 - \beta(1+\psi))(\tilde{e}^2 - e) + \beta^2 D_x D_s e,\end{aligned}$$

establishing (7.4).

To use Lemma 4, we need to bound $\tilde{e}^{-\mathrm{T}}\tilde{e}^{-1}$.

LEMMA 5. *If* $\|\tilde{e}^2 - e\| \leqslant \alpha$ *with* $\alpha \leqslant 1/3$, *then*

$$n \leqslant \tilde{e}^{-\mathrm{T}}\tilde{e}^{-1} \leqslant n + 1. \tag{7.6}$$

PROOF. For the lower bound, note that $\tilde{e}^{-\mathrm{T}}\tilde{e} = n$, while $\|\tilde{e}\| = \sqrt{n}$. Hence $\|\tilde{e}^{-1}\| \geqslant \sqrt{n}$. Now since $\tilde{e} > 0$, we certainly have $\|\tilde{e} - e\| \leqslant \alpha$ also. So

$$\|\tilde{e}^{-1} - e\| = \left\|\tilde{E}^{-1}(e - \tilde{e})\right\| \leqslant \frac{1}{1-\alpha}\|e - \tilde{e}\| \leqslant \frac{\alpha}{1-\alpha},$$

since each diagonal entry of $\tilde{E}^{-1}$ is at most $(1-\alpha)^{-1}$. Hence

$$\|\tilde{e}^{-1} - \tilde{e}\| \leqslant \alpha + \frac{\alpha}{1-\alpha} < 1$$

since $\alpha \leqslant 1/3$. Finally

$$\|\tilde{e}^{-1}\|^2 = \|\tilde{e} + (\tilde{e}^{-1} - \tilde{e})\|^2 = \|\tilde{e}\|^2 + \|\tilde{e}^{-1} - \tilde{e}\|^2 \leqslant n + 1$$

since $\tilde{e}^{\mathrm{T}}(\tilde{e}^{-1} - \tilde{e}) = n - n = 0$. Finally, we need

LEMMA 6. *If* $1 - \alpha - \beta\|d\| > 0$, *then*

$$\phi(x, s, 1) \leqslant \phi(\tilde{e}, \tilde{e}, 1) + \beta\,\nabla\phi^{\mathrm{T}} d + \frac{1}{2(1 - \alpha - \beta\|d\|)}\beta^2 d^{\mathrm{T}} d. \tag{7.7}$$

Copyright © 2001 All Rights Reserved

Downloaded from informs.org by [131.94.16.10] on 29 September 2015, at 21:31 . For personal use only, all rights reserved.

PROOF. We start with

$$\ln(1+\lambda) = \lambda - \frac{\lambda^2}{2} + \frac{\lambda^3}{3} \cdots$$

$$\geqslant \lambda - \sum_{i=2}^{\infty} \frac{|\lambda|^i}{2} = \lambda - \frac{1}{2(1-|\lambda|)}\lambda^2$$

if $|\lambda| < 1$. Now suppose $|\epsilon^2 - 1| \leqslant \alpha$ (so $|\epsilon - 1| \leqslant \alpha$ also) and $|\theta| \leqslant \|d\|$. Then

$$\ln(\epsilon + \beta\theta) = \ln \epsilon + \ln\left(1 + \beta\frac{\theta}{\epsilon}\right)$$

$$\geqslant \ln \epsilon + \beta\frac{\theta}{\epsilon} - \frac{1}{2\left(1 - \beta\left|\frac{\theta}{\epsilon}\right|\right)\epsilon^2}\beta^2\theta^2.$$

Now if $\epsilon \leqslant 1$, then

$$\left(1 - \beta\left|\frac{\theta}{\epsilon}\right|\right)\epsilon^2 = \epsilon^2 - \beta|\theta|\epsilon \geqslant \epsilon^2 - \beta|\theta|$$

$$\geqslant 1 - \alpha - \beta\|d\|,$$

while if $\epsilon > 1$, we have

$$\left(1 - \beta\left|\frac{\theta}{\epsilon}\right|\right)\epsilon^2 \geqslant 1 - \beta|\theta|\epsilon \geqslant 1 - \beta|\theta|(1+\alpha)$$

$$\geqslant 1 - \alpha - \beta\|d\|$$

since $\beta|\theta| < 1$. Hence

$$\ln(\epsilon + \beta\theta) \geqslant \ln \epsilon + \beta\frac{\theta}{\epsilon} - \frac{1}{2(1 - \alpha - \beta\|d\|)}\beta^2\theta^2. \tag{7.8}$$

Note that the first two terms are the first-order Taylor approximation of $\ln(\epsilon + \beta\theta)$. Now consider

$$\phi(x, s, 1) = \rho \ln(\tilde{e}^{\mathrm{T}}x + \tilde{e}^{\mathrm{T}}s - n) - \sum \ln x_j - \sum \ln s_j.$$

The first term is concave, hence its first-order Taylor approximation is an overestimate. For the remaining terms we use the approximation given by (7.8), where ϵ denotes a component of $\tilde{e}$ and θ a component of d_x or d_s. Since the sum of all such θ^2 is $d^{\mathrm{T}}d$, the result follows.

PROOF OF LEMMA 3. By lemmas 4 and 5, $4 \leqslant d^{\mathrm{T}}d \leqslant 5$ and so $\|d\| \leqslant 5/2$ and

$$1 - \alpha - \beta\|d\| \geqslant 1 - \frac{1}{3} - \frac{1}{15} \cdot \frac{5}{2} = \frac{1}{2}.$$

Thus each component of x and of s is positive. Also, $\beta(1 + \psi) < 1$, so that, again by

Downloaded from informs.org by [131.94.16.10] on 29 September 2015, at 21:31 . For personal use only, all rights reserved.

Copyright © 2001 All Rights Reserved

RIGHTSLINK

lemma 4,

$$\begin{aligned}(7.9)\qquad \|XSe - \zeta e\| &\leqslant (1 - \beta(1 + \psi))\|\tilde{e}^2 - e\| + \beta^2 d^{\mathrm{T}} d \\ &\leqslant (1 - \beta\psi)\alpha + \beta(\beta d^{\mathrm{T}} d - \alpha) \\ &\leqslant \alpha\zeta.\end{aligned}$$

Finally, by Lemmas 4 and 6 and the above estimates,

$$\begin{aligned}\phi(x, s, 1) &\leqslant \phi(\tilde{e}, \tilde{e}, 1) + \beta\, \nabla\phi^{\mathrm{T}} d + \frac{1}{2(1 - \alpha - \beta\|d\|)} \beta^2 d^{\mathrm{T}} d \\ &\leqslant \phi(\tilde{e}, \tilde{e}, 1) - \frac{1}{15} \cdot 2 + \frac{1}{2 \cdot (\frac{1}{2})} \left(\frac{1}{15}\right)^2 \cdot 5 \\ &= \phi(\tilde{e}, \tilde{e}, 1) - \frac{1}{9}.\end{aligned}$$

This completes the proof.

8. Discussion. We now discuss the choice of parameters made. From (5.14), ψ is about $2\rho/n$, so to achieve a good reduction in the potential function we would like $-\nabla\phi^{\mathrm{T}} d$ and hence ρ large. To maintain feasibility, we need $\beta = O(1/\|d\|)$. Let us examine the consequences of choosing ρ of order n. Then ψ is of order 1, $-\nabla\phi^{\mathrm{T}} d$ of order n, and $\|d\|$ of order $\sqrt{n}$. Hence β is of order $1/\sqrt{n}$, and the first order change in ϕ of order $\sqrt{n}$. Moreover, (7.7) shows that the error in this first-order estimate is of order 1. Hence we can achieve a decrease in ϕ of order $\sqrt{n}$ (see also Ye and Todd [32]). This would seem to imply a complexity of $O(\sqrt{n}\, L)$, since with ρ of order n we need a reduction of order n in ϕ to reduce the objective function by a constant; this bound also appears to follow from the objective function reduction in (7.3). However, such a choice of ρ does not seem to allow the algorithm to remain approximately centered, and hence the improvement cannot be sustained. While the first-order term in (7.4) is very attractive, the second-order term seems to be too large. Indeed, (7.9) shows that we require $\beta d^{\mathrm{T}} d - \alpha < 0$, and so β should be $O(1/\|d\|^2)$. This would give β of order $1/n$ and hence only a constant reduction in ϕ and a reduction in the objective function of $(1 - \gamma/n)$ for constant γ, as in Karmarkar's algorithm or Kojima et al. [15].

On the other hand, if ρ is smaller than order $\sqrt{n}$, then $-\nabla\phi^{\mathrm{T}} d$ is smaller than order 1, and we may not be able to achieve a reduction in ϕ due to second-order terms. Our choice of ρ as approximately $\sqrt{n}$ balances these requirements nicely.

We must mention a disadvantage of our choice of parameters. One of the goals of our research was to develop an algorithm that required only $\sqrt{n}\, L$ iterations as in the path-following methods and yet gave the possibility that line searches could reduce the complexity further. While line searches are possible in the path-following algorithms [12], the direction is determined directly or indirectly by seeking a reduction of the objective function by a factor $(1 - \gamma/\sqrt{n})$ for constant γ, and it is not clear that these directions are good for longer steps. Our directions arise from seeking reductions in a potential function, and one could hope that line searches would be more effective.

Unfortunately, while feasibility demands only β to be $O(1/\|d\|_\infty)$, which might be of order $\sqrt{n}$, the approximate centering places more stringent conditions. Indeed, (7.4) suggests that we must maintain $1 - \beta(1 + \psi)$ positive (or not too negative) which limits β to order 1. It may well be that the inherent curvature of the "cone" of

Downloaded from informs.org by [131.94.16.10] on 29 September 2015, at 21:31 . For personal use only, all rights reserved.

Copyright © 2001 All Rights Reserved

RIGHTSLINK

approximate centers does not allow line searches to be effective. Then either one must establish a method that successfully moves further down the path of centers, or employ a higher-order predictor of the path, as suggested by Bayer and Lagarias [5] and Adler, Karmarkar, Resende and Veiga [1]. These remarks suggest that the bidirectional search proposed by Tanabe [25] and Gonzaga [13], which amounts to varying ψ in the definition of the direction $\bar{d}$ of (5.13), may not be as successful as hoped. (As previously mentioned, the recent algorithm of Ye [31] achieves a constant reduction in our potential function without needing to remain approximately centered.)

Finally, while we hope that our framework has shown the advantages of a primal-dual setting and approximate centering in projective algorithms, the true worst-case complexity of Karmarkar's primal-only projective algorithm is still unknown. Are there examples requiring $O(nL)$ iterations, or can the improved behavior observed in practice be proved to hold in general? An interesting discussion relating to this question can be found in McDiarmid [17].

Acknowledgements. This research benefited from several very helpful conversations with Kurt Anstreicher.

References

[1] Adler, I., Karmarkar, N., Resende, M. and Veiga, G. (1986). An implementation of Karmarkar's Algorithm for Linear Programming. Report ORC86-8, Operations Research Center, University of California, Berkeley, to appear in *Math. Programming*.

[2] Anstreicher, K. (1989). The Worst-Case Step in Karmarkar's Algorithm. *Math. Oper. Res.* **14** 294–302.

[3] Barnes, E. R. (1986). A Variation on Karmarkar's Algorithm for Solving Linear Programming Problems. *Math. Programming* **36** 174–182.

[4] _____, Chopra, S. and Jensen, D. L. (1988). A Polynomial-Time Version of the Affine-Scaling Algorithm. IBM T. J. Watson Research Center, Yorktown Heights, NY.

[5] Bayer, D. and Lagarias, J. C. (1989). The Nonlinear Geometry of Linear Programming. I. Affine and Projective Scaling Trajectories. II. Legendre Transform Coordinates and Central Trajectories. *Trans. Amer. Math. Soc.* **314** 499–526, 527–581.

[6] Dikin, I. I. (1967). Iterative Solution of Problems of Linear and Quadratic Programming. *Soviet Math. Dokl.* **8** 674–675.

[7] _____, (1974). On the Speed of an Iterative Process. *Upravlyaemye Sistemi* **12** 54–60 (in Russian).

[8] de Ghellinck, G. and Vial, J.-P. (1987). An Extension of Karmarkar's Algorithm for Solving a System of Linear Homogeneous Equations on the Simplex. *Math. Programming* **39** 79–92.

[9] _____ and _____. (1986). A Polynomial Newton Method for Linear Programming. *Algorithmica* **1** 425–453.

[10] Fiacco, A. V. and McCormick, G. P. (1968). *Nonlinear Programming: Sequential Unconstrained Minimization Techniques*. John Wiley and Sons, New York.

[11] Frisch, K. R. (1955). The Logarithmic Potential Method of Convex Programming. unpublished manuscript, University Institute of Economics, Oslo, Norway.

[12] Gonzaga, C. C. (1989). An Algorithm for Solving Linear Programming Problems in $O(n^3L)$ Operations. in: *Progress in Mathematical Programming*. N. Megiddo (Ed.), Springer-Verlag, Berlin, 1–28.

[13] _____. (1987). Search Directions for Interior Linear Programming Methods. memorandum No. UCB/ERL M87/44, Electronics Research Laboratory, University of California, Berkeley.

[14] Karmarkar, N. (1984). A New Polynomial Time Algorithm for Linear Programming. *Combinatorica* **4** 373–395.

[15] Kojima, M., Mizuno, S. and Yoshise, A. (1989). A Primal-Dual Interior Point Method for Linear Programming. in: *Progress in Mathematical Programming*. N. Megiddo, (Ed.), Springer-Verlag, Berlin, 29–47.

[16] _____, _____ and _____. (1989). A Polynomial-Time Algorithm for a Class of Linear Complementarity Problems. in: *Math. Programming* **44** 1–26.

[17] McDiarmid, C. J. H. (1986). On the Improvement per Iteration in Karmarkar's Algorithm for Linear Programming. manuscript, Institute of Economics and Statistics, Oxford University, Oxford, U.K., to appear in *Math. Programming*.

[18] Megiddo, N. (1989). Pathways to the Optimal Set in Linear Programming. in: *Progress in Mathemati-*

Downloaded from informs.org by [131.94.16.10] on 29 September 2015, at 21:31 . For personal use only, all rights reserved.

Copyright © 2001 All Rights Reserved

RIGHTSLINK

cal Programming N. Megiddo (Ed.), Springer-Verlag, Berlin, 131–158.

[19] Megiddo, N. and Shub, M. (1986). Boundary Behaviour of Interior Point Algorithms in Linear Programming. Research report RJ 5319, IBM Thomas J. Watson Research Center, Yorktown Heights, NY.

[20] Mitchell, J. E. and Todd, M. J. (1986). On the relationship between the search directions in the affine and projective variants of Karmarkar's linear programming algorithm. Technical Report no. 725, School of Operations Research, Cornell University, (Ithaca, NY,

[21] Monteiro, R. C. and Adler, I. (1989). Interior Path Following Primal-Dual Algorithms. Part I: Linear Programming, *Math. Programming* **44** 27–41.

[22] _____ and _____. (1989). Interior Path Following Primal-Dual Algorithms. Part II: Convex Quadratic Programming *Math. Programming* **44** 43–66.

[23] Renegar, J. (1988). A Polynomial-Time Algorithm Based on Newton's Method for Linear Programming. *Math. Programming* **40** 59–93.

[24] Sonnevend, G. (1986). An 'Analytic Center' for Polyhedrons and New Classes of Global Algorithms for Linear (Smooth, Convex) Programming. in: *Lecture Notes in Control and Information Sciences* **84**, Springer-Verlag, New York, 866–876.

[25] Tanabe, K. (1987). Complementarity-Enforcing Centered Newton Method for Mathematical Programming. *New Methods for Linear Programming*. The Institute of Statistical Mathematics, Minatoku, Tokyo, 118–144.

[26] Todd, M. J (1988). Exploiting Special Structure in Karmarkar's Linear Programming Algorithm. *Math. Programming* **41** 97–113.

[27] _____ and Burrell, B. P. (1986). An Extension of Karmarkar's Algorithm for Linear Programming Using Dual Variables. *Algorithmica* **1** 409–424.

[28] Vaidya, P. M. (1987). An Algorithm for Linear Programming Which Requires $O(((m+n)n^2 + (m+n)^{1.5})L)$ Arithmetic Operations. preprint, AT & T Bell Laboratories, Murray Hill, NJ.

[29] Vanderbei, R. J., Meketon, M S. and Freedman, B. A. (1986). A Modification of Karmarkar's Linear Programming Algorithm. *Algorithmica* **1** 395–407.

[30] Ye, Y. (1987). Interior Algorithms for Linear, Quadratic, and Linearly Constrained Convex Programming. Ph.D. dissertation, Department of Engineering-Economic Systems, Stanford University, Stanford, CA.

[31] _____. (1986). An $O(n^3L)$ Potential Reduction Algorithm for Linear Programming. manuscript, Department of Management Sciences, The University of Iowa, Iowa City, to appear in *Math. Programming*.

[32] _____ and Todd, M. J. (1987). Containing and Shrinking Ellipsoids in the Path-Following Algorithm. manuscript, Department of Engineering-Economic Systems, Stanford University, Stanford, CA, to appear in *Math. Programming*.

SCHOOL OF OPERATIONS RESEARCH AND INDUSTRIAL ENGINEERING, COLLEGE OF ENGINEERING, CORNELL UNIVERSITY, ITHACA, NEW YORK 14853

Downloaded from informs.org by [131.94.16.10] on 29 September 2015, at 21:31 . For personal use only, all rights reserved.

Copyright © 2001 All Rights Reserved

RIGHTSLINK