Polynomial-Time Algorithms for Linear Programming Based only on Primal Scaling and Projected Gradients of a Potential Function

Robert M. Freund

Sloan W.P. No. 2048-88   August, 1988

Polynomial-Time Algorithms for Linear Programming Based only on Primal Scaling and Projected Gradients of a Potential Function

by

Robert M. Freund
Sloan School of Management
Massachusetts Institute of Technology
50 Memorial Drive
Cambridge, MA 02139

Abstract

The main result of this paper is an $O(\sqrt{n} L)$ iteration interior-point algorithm for linear programming, that does not need to follow the central trajectory to retain the complexity bound and so can be augmented by a standard line search of the potential function. This algorithm is structurally similar to Ye's $O(\sqrt{n} L)$ iteration algorithm. In the paper, we present two polynomial-time algorithms for linear programming, that use only primal affine scaling and a projected gradient of a potential function, and that do not need to follow the central trajectory, nor do they require the use of projective transformations. The first algorithm is an $O(n L)$ iteration algorithm that is an extension of Gonzaga's polynomial affine algorithm, and is based on the potential function

$$q \ln (x^T s) - \sum_{j=1}^{n} \ln (x_j)$$

where x is the vector of primal variables and s is the vector of dual slack variables, and $q = n + \sqrt{n}$. The algorithm takes either a primal step or recomputes dual variables at each iteration. The second algorithm is an $O(\sqrt{n} L)$ iteration algorithm that is an extension of the first algorithm, but uses the potential function

$$q \ln (x^T s) - \sum_{j=1}^{n} \ln (x_j) - \sum_{j=1}^{n} \ln (s_j)$$

where $q = n + \sqrt{n}$. We show that this algorithm is optimal with respect to the choice of q in the following sense. Suppose that $q = n + n^t$ where $t \geq 0$. Then the algorithm will solve the linear program in $O(n^r L)$ iterations, where $r = \max \{ t, 1 - t \}$. Thus the value of t that minimizes the complexity bound is $t = 1/2$.

Key Words: Linear program, polynomial time bound, affine scaling, interior method.

1. Introduction.

The current interest in interior methods for linear programming and its extensions stems from the seminal work of Karmarkar [8] , who presented an algorithm for linear programming that requires at most $O(n\,L)$ iterations. The work per iteration involves solving a least-squares problem, which requires $O(n^3)$ operations; however, by solving the least-squares problem inexactly in $O(n^2)$ operations by performing rank-1 updates, Karmarkar gave a clever argument that reduces the overall complexity bound to $O(n^{3.5}\,L)$ iterations. Todd and Burrell [13] suggested that Karmarkar's algorithm could be implemented with a standard line-search at each iteration, with a substantial potential for decreasing the actual iteration count, at the cost of an increase in the complexity bound by $\sqrt{n}$ to $O(n^4\,L)$ operations. Anstreicher [2] has shown a safeguarded line-search that retains the overall complexity bound of $O(n^{3.5}\,L)$ iterations , but it is not clear whether this method is more or less efficient in practice than a standard line-search. Karmarkar's algorithm and all of its projective transformation variants and extensions (Todd and Burrell [13], Anstreicher [1], Gay [5], Rinaldi [2], Ye [16], and [4], for example) all retain the $O(n\,L)$ iteration count. Gonzaga [7] has presented an $O(n\,L)$ iteration count algorithm that uses a potential function but does not perform projective transformations.

In [11], Renegar was the first to develop an interior method for solving a linear program with an $O(\sqrt{n}\,L)$ iteration count. The algorithm works by tracing the central trajectory (see Megiddo [9] and Bayer and Lagarias [3]) with Newton steps. Other central trajectory path-following algorithms with the $O(\sqrt{n}\,L)$ iteration count have been developed since then, see Gonzaga [6], Monteiro and Adler [10], Vaidya [15], and Todd and Ye [14], among others. If the initial feasible solution to the linear program is not near the central trajectory, the problem is artificially augmented at the start so that the initial solution is near the central trajectory, and so the algorithm can be initiated. Because the iterates must stay close to the central trajectory, the use of a line-search does not appear as promising for increasing the performance, as it does for projective transformation based algorithms. Nevertheless, the worst-case complexity bound on the iteration count for a central-trajectory path-following algorithm is a $\sqrt{n}$ improvement over the bound for a projective transformation algorithm.

Quite recently, Ye [17, 18] has presented an interior-point algorithm that has the advantage of the central trajectory methods (an $O(\sqrt{n}\,L)$ iteration count) along with the advantages of the projective transformation methods (the method can be initiated directly from any interior solution and the use of a standard line-search appears promising).

In this paper, we present two interior-point algorithms for linear programming, that use only primal affine scaling and a projected gradient of a potential function. The second algorithm, which is structurally similar to Ye's algorithm [18], has an $O(\sqrt{n} L)$ iteration count, and shares the advantages of starting the algorithm directly from any interior point, as well as being able to use a line-search at each iteration.

The first algorithm, presented in Section 3, is an extension of Gonzaga's algorithm for linear programming, but does not make the restrictive assumption of a known optimal objective function value. Performance of the algorithm is measured with the potential function

$$F(x, s) = q \ln (x^T s) - \sum_{j=1}^{n} \ln(x_j)$$

where x is the vector of primal variables and s is the vector of dual slack variables, and $q = n + \sqrt{n}$. At each iteration, the algorithm either takes a primal step or recomputes the dual feasible solution. Primal iterates decrease the potential function by at least .18, whereas dual iterates decrease the potential function by at least $.2\sqrt{n}$. This leads to an overall iteration bound of $O(n L)$ iterations.

In Section 4, we modify the algorithm of Section 3, and measure its performance with the primal-dual potential function

$$G(x, s) = q \ln (x^T s) - \sum_{j=1}^{n} \ln (x_j) - \sum_{j=1}^{n} \ln (s_j)$$

where again $q = n + \sqrt{n}$. Like the algorithm of Section 3, at each iteration this algorithm either takes a primal step or recomputes the dual feasible solution. Both primal and dual iterates decrease the potential function by at least .02, and this leads to an overall iteration bound of $O(\sqrt{n} L)$ iterations.

The analysis of Section 4 suggests that the factor of $\sqrt{n}$ plays a very important role in the potential function parameter $q = n + \sqrt{n}$. This is examined from a complexity point of view in Section 5, where it is shown that $q = n + \sqrt{n}$ is optimal in the following sense. Suppose that $q = n + n^t$, where $t \geq 0$. Then the algorithm will solve the linear program

in $O(n^r L)$ iterations, where $r = \max\{t, 1 - t\}$. Thus the the value of t that minimizes the complexity bound is $t = 1/2$, which yields the $O(\sqrt{n} L)$ iterations algorithm of Section 4.

Section 5 contains concluding remarks.

2. Notation, Assumptions, and Preliminaries.

If x or s is a vector in R^n, then X or S refers to the $n \times n$ diagonal matrix with diagonal entries corresponding to the components of x or s. Let e be the vector of ones, $e = (1, 1, \ldots, 1)$, where typically the dimension of e is n. If $x \in R^n$, then $\|x\|_p$ refers to the p-norm of x, where $1 \le p \le \infty$.

Our concern is with solving a linear program of the form

LP: minimize $c^T x$

          s.t. $Ax = b$

                  $x \ge 0$,

where for convenience, we assume that A is $m \times n$ with rank m. We assume (i) that LP has a strictly positive feasible solution $\bar{x} > 0$, and (ii) that the set of optimal solutions of LP is nonempty and bounded. The dual of LP is given by

LD: minimize $b^T \pi$

          s.t. $A^T \pi + s = c$

                  $s \ge 0$.

The second assumption above is equivalent, via a theorem of the alternative, to the assumption that there exists a solution $(\bar{\pi}, \bar{s})$ to LD with $\bar{s} > 0$.

Let $\bar{x}$ be a strictly positive feasible solution to LP , and consider the rescaled primal and dual pair:

$$\text{LP}\bar{X}: \quad \text{minimize} \quad c^T \bar{X} z \qquad\qquad \text{LD}\bar{X}: \quad \text{maximize} \quad b^t\pi$$

$$\text{s.t.} \quad A\bar{X} z = b \qquad\qquad\qquad \text{s.t.} \quad \bar{X} A^T\pi + t = \bar{X}c$$

$$z \geq 0 \qquad\qquad\qquad\qquad\qquad t \geq 0 .$$

Note that x is feasible for LP if and only if $z = \bar{X}^{-1}x$ is feasible for LP$\bar{X}$, and that (π, s) is feasible for LD if and only if $(\pi, t) = (\pi, \bar{X}s)$ is feasible for LD$\bar{X}$. If $\bar{x}$ and $(\bar{\pi}, \bar{s})$ are primal and dual feasible, then the duality gap between $c^T\bar{x}$ and $\bar{\pi}^T b$ is given by $c^T\bar{x} - \bar{\pi}^T b = \bar{x}^T\bar{s}$. We will consider two types of potential functions for LP . The first is a $\underline{\text{primal potential function}}$, given by

$$F(x, s) = q \ln (x^T s) - \sum_{j=1}^{n} \ln (x_j) \tag{2.1}$$

where $x \in P = \{x \in R^n | Ax = b, x \geq 0\}$ and $s \in D = \{s \in R^n | s \geq 0 \text{ and } A^T\pi + s = c \text{ for some } \pi \in R^m\}$ and q is a given parameter. Note that the dual slack variables enter $F(x,s)$ through the duality gap term only. The $\underline{\text{primal-dual potential function}}$ is given by

$$G(x, s) = q \ln (x^T s) - \sum_{j=1}^{n} \ln (x_j) - \sum_{j=1}^{n} \ln (s_j) . \tag{2.2}$$

It is elementary to verify that scaling LP to LP$\bar{X}$ translates the primal potential function $F(x, s)$ by an additive constant

$$\sum_{j=1}^{n} \ln (\bar{x}_j) ,$$

and that this scaling leaves the primal-dual potential function $G(x, s)$ unchanged.

Suppose L is the bit-size of a given instance of LP , and suppose we have an algorithm for LP that reduces the primal potential function F(x, s) by an amount greater than or equal to δ at each iteration, where $\delta > 0$ is a fixed quantity. Then as in Gonzaga [7] , the algorithm can be used to solve LP in $O((q/\delta) L)$ iterations, as long as the initial values (x^0, s^0) of (x, s) satisfy $F(x^0, s^0) \leq O((q/\delta)L)$. This is because in order to round to an optimal solution to LP , we need $(x^T s) \leq 2^{-L}$, and the bound on the reduction of $\ln(x^T s)$ through the potential function F(x, s) is proportional to δ , and inversely proportional to q . However, if instead we use the primal-dual function G(x, s) we have:

<u>Proposition 2.1.</u> (see also Ye [18]): If an algorithm for solving LP reduces the primal-dual potential function G(x, s) by an amount greater than or equal to δ at each iteration, where $\delta > 0$ is a fixed quantity, then the algorithm can be used to solve LP in $O[((q-n)/\delta) L]$ iterations, so long as the initial feasible primal-dual values (x^0, s^0) satisfy $G(x^0, s^0) \leq O[((q-n)/\delta) L]$.

The proof of Proposition of 2.1 follows from the inequality in the following remark:

<u>Remark 2.1.</u> If $x > 0$ and $s > 0$ are vectors in R^n , then

$$n \ln(x^T s) - \sum_{j=1}^{n} \ln(x_j) - \sum_{j=1}^{n} \ln(s_j) \geq n \ln(n). \tag{2.3}$$

<u>Proof.</u> Let $t = Xs$. Then

$$n \ln(x^T s) - \sum_{j=1}^{n} \ln(x_j) - \sum_{j=1}^{n} \ln(s_j) = n \ln(e^T t) - \sum_{j=1}^{n} \ln(t_j) .$$

However,

$$\prod_{j=1}^{n} (t_j/e^T t) \leq (1/n)^n$$

because the vector $t = e$ is a maximizer of the product term. Thus,

$$\sum_{j=1}^{n} \ln(t_j / e^T t) \leq -\ln(n) \quad,$$

whereby,

$$n \ln(n) \leq n \ln(e^T t) - \sum_{j=1}^{n} \ln(t_j) = n \ln(x^T s) - \sum_{j=1}^{n} \ln(x_j) - \sum_{j=1}^{n} \ln(s_j) . \quad \blacksquare$$

Proof of Proposition 2.1. After $O(L(q-n)/\delta)$ iterations of the algorithm, the current primal-dual iterates $(\bar{x}, \bar{s})$ satisfy

$$G(\bar{x}, \bar{s}) \leq -(q-n) L ,$$

and so

$$q \ln(\bar{x}^T \bar{s}) - \sum_{j=1}^{n} \ln(\bar{x}_j) - \sum_{j=1}^{n} \ln(\bar{s}_j) \leq -(q-n) L .$$

However, by combining the above with inequality (2.3), we obtain

$$(q-n) \ln(\bar{x}^T \bar{s}) \leq -(q-n) L - n \ln(n) ,$$

and so $\ln(\bar{x}^T \bar{s}) \leq -L$,

whereby $\bar{x}$ and $\bar{s}$ can be rounded to optimal primal and dual solutions. $\quad \blacksquare$

3. A Primal Potential Function Algorithm for LP that Converges in O(n L) Iterations

In this section we consider an algorithm for solving LP by using the primal potential function F(x, s) of (2.1). Our principal motivation for this algorithm is the work of Gonzaga [7], who presented an O(n L) iteration algorithm based on a primal potential function where the optimal objective function value is presumed known. Herein, we remove this restrictive assumption, and demonstrate a primal-dual property underlying Gonzaga's methodology that will be used in primal-dual potential function based algorithms in Sections 4 and 5.

Let us assume that $\bar{x}$ and $(\bar{\pi}, \bar{s})$ are primal and dual feasible solutions, and that $\bar{x} > 0$. Because scaling does not affect the potential function, we assume now that $\bar{x} = e$, and so the current duality gap is $e^T \bar{s}$. Let us now compute the projection of $\nabla_x F(e, \bar{s})$, the gradient of F(x, s) at $(e, \bar{s})$ in the x-coordinates, onto the null space of A . The gradient of F(x, s) in the x-coordinates at $(e, \bar{s})$ is given by:

$$g = \nabla_x F(e, \bar{s}) = \left(\frac{q}{e^T \bar{s}} \right) \bar{s} - e \qquad (3.0)$$

and its projection onto the null-space of A is given by

$$\left[I - A^T (AA^T)^{-1} A \right] \left[\left(\frac{q}{e^T \bar{s}} \right) \bar{s} - e \right] \qquad .$$

If we replace $(e^T \bar{s})$ in the above expression by the unknown quantity Δ , we obtain the direction function

$$d(\Delta) = \left[I - A^T (AA^T)^{-1} A \right] \left[\left(\frac{q}{\Delta} \right) \bar{s} - e \right] \qquad (3.1)$$

where Δ is a positive scalar.

From (3.0) and (3.1), the projected gradient is $d(e^{T}s)$. Gonzaga's algorithm [7] is motivated by demonstrating that $||d(e^{T}s)||_2 \geq 1$. Herein, we will alter this condition to analyze the consequences of the condition $||d(e^{T}s)||_2 \geq .8$, where the number .8 was fairly arbitrarily chosen, and in fact any fixed scalar in the range $(0, 1)$ would do. If $d(e^{T}s)$ satisfies $||d(e^{T}s)||_2 \geq .8$, then the potential function $F(x, s)$ can be reduced by at least a constant amount of .18 by taking a step in the direction $-d(e^{T}s)$, from the current point $\overline{x} = e$, as the next Proposition shows.

Proposition 3.1. Let $\overline{d} = d(e^{T}s)$. If $||\overline{d}||_2 \geq .8$, then

$$F(e - .38\overline{d}/||\overline{d}||_2, \overline{s}) \leq F(e, \overline{s}) - .18 \quad .$$

Proof. Let us consider $F(e - \alpha\overline{d}/||\overline{d}||_2, \overline{s})$, where $0 \leq \alpha < 1$.

Then

$$F(e - \alpha\overline{d}/||\overline{d}||_2, \overline{s}) - F(e, \overline{s})$$

$$= q \ln ((e^{T}s) - \alpha\overline{s}^{T}\overline{d}/||\overline{d}||_2) - \sum_{j=1}^{n} \ln (1 - \alpha\overline{d}_j/||\overline{d}||_2) - q \ln (e^{T}s)$$

$$= q \ln\left(\frac{\alpha\overline{s}^{T}\overline{d}}{(e^{T}s)||\overline{d}||_2}\right) - \sum_{j=1}^{n} \ln\left(1 - \frac{\alpha\overline{d}_j}{||\overline{d}||_2}\right) .$$

By the local properties of the logarithm function given in the Appendix in Proposition A.1., we have:

$$F(e - \alpha\overline{d}/||\overline{d}||_2, \overline{s}) - F(e, \overline{s}) \leq \left(\frac{-q\alpha\overline{s}^{T}\overline{d}}{(e^{T}s)||\overline{d}||_2}\right) + \sum_{j=1}^{n} \alpha\overline{d}_j/||\overline{d}||_2 + \sum_{j=1}^{n} \frac{(\alpha\overline{d}_j/||\overline{d}||_2)^2}{2(1-\alpha)}$$

$$= \frac{\alpha}{||\overline{d}||_2}\left[-\frac{q}{(e^{T}s)}\overline{s} + e\right]^{T}\overline{d} + \frac{\alpha^2}{2(1-\alpha)}$$

$$= \frac{-\alpha g^T \overline{d}}{||\overline{d}||_2} + \frac{\alpha^2}{2(1-\alpha)}$$

where g is the gradient of $F(x, s)$ at $(e, \overline{s})$ in the x-coordinates, see (3.0). However, noting that $g^T\overline{d} = \overline{d}^T\overline{d} = ||\overline{d}||_2^2$, this last expression becomes:

$$F(e - \alpha\overline{d} /||\overline{d}||_2, \overline{s}) - F(e, \overline{s}) \leq -\alpha||\overline{d}||_2 + \frac{(\alpha)^2}{2(1-\alpha)} . \tag{3.2}$$

Upon setting $\alpha = .38$ and noting that $||\overline{d}||_2 = .8$, we obtain

$$F(e - \alpha\overline{d} /||\overline{d}||_2, \overline{s}) - F(e, \overline{s}) \leq .18 . \qquad \blacksquare$$

In order to analyze the case when $||\overline{d}||_2 < .8$, we proceed as follows.

Noting that

$$d(\Delta) = \left[I - A^T(AA^T)^{-1}A \right] \left(\left(\frac{q}{\Delta} \right)c - e \right),$$

we expand and rewrite this expression as

$$A^T(AA^T)^{-1}A \left(c - \frac{\Delta}{q}e \right) + \left(\frac{\Delta}{q} \right)(e + d(\Delta)) = c \tag{3.3}$$

and define

$$\pi(\Delta) = (AA^T)^{-1}A \left(c - \frac{\Delta}{q} e \right) \tag{3.4}$$

and

$$s(\Delta) = \left(\frac{\Delta}{q} \right)(e + d(\Delta)) . \tag{3.5}$$

Then $\pi(\Delta)$ and $s(\Delta)$ satisfy $A^T\pi(\Delta) + s(\Delta) = c$ and so are dual feasible if $s(\Delta) \geq 0$. Furthermore, from (3.5), $s(\Delta) \geq 0$ if $\|d(\Delta)\|_2 \leq 1$ and $\Delta > 0$. Thus, if $\|d(e^{T\bar{s}})\|_2 < .8$, then $s(e^{T\bar{s}}) \geq 0$ and $\pi(e^{T\bar{s}})$, $s(e^{T\bar{s}})$ are dual feasible. Next note from (3.1) that $\lim_{\Delta \to 0} \|d(\Delta)\|_2 = +\infty$, unless $[I - A^T(AA^T)^{-1}A]\, c = 0$, in which case c^Tx is

constant on the feasible region of LP and $\bar{x} = e$ solves LP. Thus, by the continuity of $d(\Delta)$ in the range $\Delta \in (0, \infty)$, if $\|d(e^{T\bar{s}})\|_2 < .8$, there exists a value $\tilde{\Delta} \in (0, e^{T\bar{s}})$ such that $\|d(\tilde{\Delta})\|_2 = .8$. If we set $\tilde{\pi} = \pi(\tilde{\Delta})$ and $\tilde{s} = s(\tilde{\Delta})$, we then obtain a $.2\sqrt{n}$

decrease in $F(e, s)$ if $q \geq n + \sqrt{n}$, as the next Lemma demonstrates.

__Lemma 3.1. (Change in Dual Variables).__ Suppose $q \geq n + \sqrt{n}$. Suppose $\|d(e^{T\bar{s}})\|_2 < .8$, and that c^Tx is not constant on the feasible region of LP. Let $\tilde{\Delta} \in (0, e^{T\bar{s}})$ satisfy $\|d(\tilde{\Delta})\|_2 = .8$ and let $\tilde{\pi} = \pi(\tilde{\Delta})$ and $\tilde{s} = s(\tilde{\Delta})$. Then $\tilde{\pi}, \tilde{s}$ are feasible for LD and

$$F(e, \tilde{s}) - F(e, \bar{s}) \leq - .2\sqrt{n}.$$

__Proof:__ Note that $e^T\tilde{s} = e^Ts(\tilde{\Delta}) = \dfrac{\tilde{\Delta}}{q}[e^Te + e^Td(\tilde{\Delta})]$, from (3.5).

However, $e^Te = n$, and $e^Td(\tilde{\Delta}) \leq \|d(\tilde{\Delta})\|_1 \leq \sqrt{n}\,\|d(\tilde{\Delta})\|_2 = .8\sqrt{n}$, so that

$$e^T\tilde{s} \leq \frac{\tilde{\Delta}}{q}(n + .8\sqrt{n}) \leq \frac{e^T\bar{s}}{q}(n + .8\sqrt{n}). \text{ Therefore } \left(\frac{e^T\tilde{s}}{e^T\bar{s}}\right) \leq \left(\frac{n + .8\sqrt{n}}{q}\right).$$

Next note that $F(e, \tilde{s}) - F(e, \bar{s}) = q \ln \left(\dfrac{e^T\tilde{s}}{e^T\bar{s}}\right)$

$$\leq q \ln\left(\frac{n + .8\sqrt{n}}{q}\right) = q \ln\left(1 - \frac{.2\sqrt{n}}{q}\right) \leq - .2\sqrt{n}. \qquad \blacksquare$$

Summarizing Proposition 3.1 and Lemma 3.1, we have:

<u>Remark 3.1</u>: Assume $q = n + \sqrt{n}$. Then

a) if $\|d(e^{T}\bar{s})\|_2 \geq .8$, we can decrease $F(x, s)$ by $.18$ by taking a step in the primal in the direction $-d(e^{T}\bar{s})$.

b) if $\|d(e^{T}\bar{s})\|_2 < .8$, we can decrease $F(x, s)$ by $.2\sqrt{n}$ by replacing $\bar{s}$ by $\tilde{s} = s(\tilde{\Delta})$ where $\tilde{\Delta} \in (0, e^{T}\bar{s})$, and $\|d(\tilde{\Delta})\|_2 = .8$.

This leads to the following algorithm for solving LP . Let (A, b, c) be the data for the LP , let $\bar{x}$ and $(\bar{\pi}, \bar{s})$ be initial primal and dual solutions, where $\bar{x} > 0$, and let $\mu > 0$ be the optimality tolerance. We have:

<u>Algorithm 1.</u> $(A, b, c, \bar{x}, \bar{s}, \bar{\pi}, \mu)$

<u>Step 0.   Initialize.</u>
$$\text{Set } q = n + \sqrt{n}, \gamma = .8 \ .$$

<u>Step 1.   Test for Optimality.</u>

$$\text{If } \bar{x}^{T}\bar{s} \leq \mu, \text{ stop.}$$

<u>Step 2.   Rescale.</u>

$$\text{Set } \bar{A} = A\bar{X}, \bar{c} = \bar{X}c, \bar{t} = \bar{X}\bar{s} \ .$$

<u>Step 3.   Compute Direction Function.</u>

$$d(\Delta) = [I - \bar{A}(\bar{A}\bar{A}^{T})^{-1}\bar{A}] \left[\left(\frac{q}{\Delta} \right) \bar{t} - e \right]$$

If $\|d(e^{T}\bar{t})\|_2 \geq \gamma$, go to Step 5. Otherwise go to Step 4.

Step 4. Recompute Dual Variables.

Solve for $\tilde{\Delta} \in (0, e^{T}\tilde{t})$ such that $\tilde{\Delta}$ is the smallest value of Δ for which $\|d(\tilde{\Delta})\|_2 = \gamma$.

Set $\tilde{t} = \left(\dfrac{\tilde{\Delta}}{q}\right)(e + d(\tilde{\Delta}))$

and $\tilde{\pi} = (\bar{A}\bar{A}^{T})^{-1}\bar{A}\left[\bar{c} - \left(\dfrac{\tilde{\Delta}}{q}\right)e\right]$

Set $\bar{\pi} = \tilde{\pi}$ and $\bar{s} = \bar{X}^{-1}\tilde{t}$, and go to Step 1.

Step 5. Take Step in Primal Variable Space.

Set $\bar{d} = d(e^{T}\bar{t})$. Set $\tilde{z} = e - .38\,\bar{d}/\|\bar{d}\|_2$.

Set $\bar{x} = \bar{X}\tilde{z}$, and return to Step 1.

Note that in Algorithm 1, that at Step 2 $\bar{t}$ is the scaled version of the current dual slack $\bar{s}$; $(\bar{\pi}, \bar{s})$ is feasible in LD , and $(\bar{\pi}, \bar{t})$ is feasible in $LD\bar{X}$, and e is feasible in $LP\bar{X}$. At the end of Steps 4 and 5, respectively, the dual and primal variables are then rescaled to the original problems LD and LP , respectively. Note also that the value of $\gamma = .8$ was chosen somewhat arbitrarily; instead, one could replace .8 by any constant $\gamma \in (0, 1)$. The step length in Step 5 would have to be adjusted, however.

Next, note that after the algorithm leaves Step 4, the next direction function $d(\Delta)$ will be the same as the current direction function. Thus, there is no need to re-solve the least squares problem in Step 3 at the next iteration. Furthermore, at the next iterate, we will have $\|d(e^{T}\bar{t})\|_2 \geq .8$ in Step 3, so we could just go straight to Step 5 directly.

Note that at Step 5, we could replace $\tilde{z} = e - .38\,\bar{d}/\|\bar{d}\|_2$ by $\tilde{z} = e - \alpha\bar{d}$ where α is determined by a line-search that (inexactly) minimizes the potential function $F(e - \alpha\bar{d}, \bar{t})$ over the range $\alpha \in (.38/\|\bar{d}\|_2, \infty)$. Also, the choice of $\tilde{\Delta}$ in step 4 could be chosen by a line-search of Δ with the potential function $F(e, \dfrac{\tilde{\Delta}}{q}[e + d(\Delta)])$ where $\Delta \in (0, e^{T}\bar{t})$.

Finally, note that the algorithm need not be started with a dual feasible solution. In fact, it need only be started by a lower bound B on the primal objective value. One would then replace the expression for $d(\Delta)$ in Step 3 by

$$d(\Delta) = [I - \bar{A}^T(\bar{A}\,\bar{A}^T)^{-1}\bar{A}]\left[\left(\frac{q}{\Delta}\right)\bar{c} - e\right],$$

as the two expressions are equivalent. Also, one would replace $e^T\bar{t}$ by $c^T\bar{x} - B$ in Steps 3, 4, and 5. The algorithm must eventually compute feasible dual variables, unless B is the optimal value of LP.

Finally, note that solving for $\tilde{\Delta}$ in Step 4 is equivalent to solving a quadratic form in the quantity $\left(\frac{1}{\Delta}\right)$ and so can be done exactly (within roundoff error) by using the quadratic formula. This is reminiscent of the dual-variable update procedure presented in Anstreicher [1], that also solves for a vector of appropriate norm in order to update dual variables.

From Remark 3.1, and the invariance of the potential function $F(x, s)$ (modulo an additive constant) under scaling, we have:

Lemma 3.2. If (x^0, s^0) are the initial values of $\bar{x}$ and $\bar{s}$ in Algorithm 1 and $F(x^0, s^0) \leq O(n\,L)$, then Algorithm 1 solves LP in $O(n\,L)$ iterations.

Proof: From Remark 3.1, each iteration reduces the potential function $F(x,s)$ by at least $\delta = .18$, and $q = n + \sqrt{n} \leq 2n$, so $q = O(n)$. Thus after $O(n\,L)$ iterations, Algorithm 1 solves LP. ∎

4. A Primal-Dual Potential Function Algorithm that solves LP in $O(\sqrt{n}L)$ Iterations.

In this section, we modify the ideas and the algorithm presented in Section 3, by using the primal-dual potential function $G(x, s)$ of (2.2) with $q = n + \sqrt{n}$. Our principal motivation for this algorithm is again the work of Gonzaga [7]. The resulting algorithm converges in $O(\sqrt{n}\,L)$ iterations, and is an algorithm similar to the $O(\sqrt{n}\,L)$ iteration algorithm of Ye [18]. There is a strong connection between the algorithm of this section and Ye's algorithm, which is discussed at the end of the section.

We begin the analysis by returning to the analysis developed at the start of Section 3. We assume that $\bar{x}$ and $(\bar{\pi}, \bar{s})$ are our primal and dual feasible solutions, that $\bar{x} > 0$ and that the LP has been rescaled so that $\bar{x} = e$, so that the current duality gap is $e^T \bar{s}$. We compute the direction function $d(\Delta)$ as in (3.1), and compute $\bar{d} = d(e^T \bar{s})$ which is the projected gradient of $G(x, s)$ in the x-coordinates, at $(x, s) = (e, \bar{s})$. Then if $\|d(e^T \bar{s})\|_2 \geq .22$, we can achieve a constant decrease in the potential function $G(x, s)$ by taking a step in the direction $-d(e^T \bar{s})$ from the current primal feasible point e, as Proposition 4.1 shows.

__Proposition 4.1.__ Let $\bar{d} = d(e^T \bar{s})$. Then if $\|\bar{d}\|_2 \geq .22$,

$$G(e - (1/6)\bar{d}/ \|\bar{d}\|_2, \ \bar{s}) \ \leq \ G(e, \bar{s}) - .02 \ .$$

__Proof:__ The proof is exactly the same as for Proposition 3.1. At the end of the proof we use $\alpha = 1/6$ to obtain the result. ∎

If $\|d(e^T \bar{s})\|_2 < .22$, however, we seek to modify our dual solution. Note that at $\Delta = e^T \bar{s}$,

$$s(\Delta) = \left(\frac{\Delta}{q}\right)(e + d(\Delta))$$

is "close" to a scaled version of the vector e, because $\|d(\Delta)\|_2 < .22$. Thus, $s(\Delta)$ is nicely centered, as the next two Lemmas demonstrate.

__Lemma 4.2.__ If $s = \left(\frac{\Delta}{q}\right)((e + d))$, where $\|d\|_2 < .22$, then $\| s - \left(\frac{e^T s}{n}\right) e \|_2 \ \leq \ .565 \left(\frac{e^T s}{n}\right)$.

__Proof.__ If $s = \left(\frac{\Delta}{q}\right)((e + d))$, then

$$e^T s \ = \ \frac{\Delta}{q}(n + e^T d) \ \geq \ \frac{\Delta}{q}(n - \|d\|_1) \ \geq \ \frac{\Delta}{q}(n - \sqrt{n}\|d\|_2)$$

$$\geq \ \frac{\Delta}{q}(n - 22\sqrt{n}) \ \geq \ \frac{\Delta}{q}(.78n) \ .$$

Thus,

$$\frac{\Delta}{q} \le \left(\frac{e^T s}{.78n}\right) .$$

(4.1)

Next note that

$$s - \left(\frac{e^T s}{n}\right) e = \frac{\Delta}{q} (e + d) - \left(\frac{\Delta}{q}\right) \left(\frac{n + e^T d}{n}\right) e = \frac{\Delta}{q} \left[d - \left(\frac{e^T d}{n}\right) e\right] .$$

Thus,

$$\| s - \left(\frac{e^T s}{n}\right) e \|_2 \le \left[\|d\|_2 + \frac{\|d\|_1}{\sqrt{n}}\right] \le \frac{\Delta}{q} \left[\|d\|_2 + \|d\|_2\right] \le \frac{.44\Delta}{q} .$$

(4.2)

Combining (4.1) and (4.2) yields the desired result. ∎

Lemma 4.3. If $s = \frac{\Delta}{q} (e + d)$, where $\|d\|_2 < .22$, then

$$\sum_{j=1}^{n} \ln (s_j) \ge n \ln\left(\frac{e^T s}{n}\right) - .367.$$

Proof:

Let $r = \left(\frac{n}{e^T s}\right) s$. Then $e^T r = n$ and $\|r - e\|_2 = \left(\frac{n}{e^T s}\right) \|s - \left(\frac{e^T s}{n}\right) e\|_2 \le .565$,

from Lemma 4.2.

Thus, from Proposition A.1 of the Appendix,

$$\sum_{j=1}^{n} \ln(r_j) \ge \sum_{j=1}^{n} (r_j - 1) - \sum_{j=1}^{n} \frac{(r_j - 1)^2}{2(1 - .565)} = -\frac{\|r - e\|_2^2}{.87} = -\frac{(.565)^2}{.87} \ge -.367 .$$

Next note that

$$\sum_{j=1}^{n} \ln(s_j) = \sum_{j=1}^{n} \ln(r_j) + n \ln\left(\frac{e^T s}{n}\right) \geq n \ln\left(\frac{e^T s}{n}\right) - .367. \quad \blacksquare$$

Suppose that $\|d(e^T \bar{s})\|_2 < .22$. Just as in Section 3 (but with .8 replaced by .22), we seek a value of $\tilde{\Delta} \in (0, e^T \bar{s})$ for which $\|d(\tilde{\Delta})\|_2 = .22$. Because $\lim\limits_{\Delta \to 0} \|d(\Delta)\|_2 = +\infty$ unless $c^T x$ is constant on the feasible region of LP (and so $\bar{x} = e$ solves LP), then such a value $\tilde{\Delta}$ must indeed exist. We then set $\tilde{s} = s(\tilde{\Delta})$ and $\tilde{\pi} = \pi(\tilde{\Delta})$ where $s(\Delta)$ and $\pi(\Delta)$ are given by (3.4) and (3.5). $\tilde{s}$ and $\tilde{\pi}$ are feasible for the dual LD . The next Lemma demonstrates that replacing $\bar{s}$ by $\tilde{s}$ results in a fixed decrease in the potential function $G(x, s)$ of at least .02, if $q = n + \sqrt{n}$.

Lemma 4.4 (Potential Function Decrease From Change of Dual Variables)

Suppose $q = n + \sqrt{n}$, and $\|d(e^T \bar{s})\|_2 < .22.$, and that $c^T x$ is not constant on the feasible region of LP . Let $\tilde{\Delta} \in (0, e^T \bar{s})$ satisfy $\|d(\tilde{\Delta})\|_2 = .22$ and let $\tilde{s} = s(\tilde{\Delta})$ and $\tilde{\pi} = \pi(\tilde{\Delta})$. Then, $(\tilde{\pi}, \tilde{s})$ are feasible for LD, and

$$G(e, \tilde{s}) - G(e, \bar{s}) \leq - .02.$$

Proof: As in Section 3, it is straightforward to check that $A^T \tilde{\pi} + \tilde{s} = c$, and $\|d(\tilde{\Delta})\|_2 = .22$ implies that $\tilde{s} > 0$, so $(\tilde{\pi}, \tilde{s})$ are dual feasible. Now we compute

$$G(e, \tilde{s}) - G(e, \bar{s}) = q \ln(e^T \tilde{s}) - \sum_{j=1}^{n} \ln(\tilde{s}_j) - q \ln(e^T \bar{s}) + \sum_{j=1}^{n} \ln(\bar{s}_j)$$

$$= q \ln\left(\frac{e^T \tilde{s}}{e^T \bar{s}}\right) - \sum_{j=1}^{n} \ln(\tilde{s}_j) + \sum_{j=1}^{n} \ln(\bar{s}_j). \qquad (4.3)$$

However, from Lemma 4.3, we have

$$\sum_{j=1}^{n} \ln(\tilde{s}_j) \geq n \ln\left(\frac{e^{T}\tilde{s}}{n}\right) - .367 .$$ (4.4)

Also, $$\sum_{j=1}^{n} \ln(\bar{s}_j) \leq \max\left\{\sum_{j=1}^{n} \ln(s_j) \mid s > 0, \ e^{T}s = e^{T}\bar{s}\right\} = n \ln\left(\frac{e^{T}\bar{s}}{n}\right).$$ (4.5)

Combining (4.3), (4.4) and (4.5) yields

$$G(e, \tilde{s}) - G(e, \bar{s}) \leq q \ln\left(\frac{e^{T}\tilde{s}}{e^{T}\bar{s}}\right) - n \ln\left(\frac{e^{T}\tilde{s}}{n}\right) + .367 + n \ln\left(\frac{e^{T}\bar{s}}{n}\right)$$

$$= (q - n) \ln\left(\frac{e^{T}\tilde{s}}{e^{T}\bar{s}}\right) + .367$$

However,

$$e^{T}\tilde{s} = \left(\frac{\tilde{\Delta}}{q}\right)(n + e^{T}d(\tilde{\Delta})) \leq \left(\frac{e^{T}\bar{s}}{q}\right)(n + \sqrt{n}\|d(\tilde{\Delta})\|_2) = \left(\frac{e^{T}\bar{s}}{q}\right)(n + .22\sqrt{n}) .$$

Thus, $$\left(\frac{e^{T}\tilde{s}}{e^{T}\bar{s}}\right) \leq \frac{n + .22\sqrt{n}}{q} = 1 - \frac{q - n - .22\sqrt{n}}{q} .$$ (4.6)

Then

$$G(e, \tilde{s}) - G(e, \bar{s}) \leq (q - n) \ln\left(\frac{e^{T}\tilde{s}}{e^{T}\bar{s}}\right) + .367$$

$$\leq \frac{-(q - n)(q - n - .22\sqrt{n})}{q} + .367$$ (4.7)

$$= -\frac{.78n}{n + \sqrt{n}} + .367 \leq -\frac{.78n}{2n} + .367 \leq - .02. \blacksquare$$

Intuitively, the proof of Lemma 4.4 works for two reasons. First, $\tilde{s}$ is "close" to the

vector $\left(\dfrac{e^{T}\tilde{s}}{n}\right)e$, as shown in Lemmas 4.2 and 4.3. Thus, the barrier function $\displaystyle\sum_{j=1}^{n} \ln(\tilde{s}_j)$ is

well-controlled. Second, by requiring that we adjust $\tilde{\Delta}$ downward from $e^{T}\tilde{s}$ to a point

where $\|d(\tilde{\Delta})\|_2 = .22$, we guarantee that $e^{T}\tilde{s}$ is sufficiently less than $e^{T}\bar{s}$.

Summarizing Proposition 4.1 and Lemma 4.4, we have:

<u>Remark 4.1</u>. Suppose $q = n + \sqrt{n}$. Then

    a) If $\|d(e^{T}\bar{s})\|_2 \geq .22$, we can decrease $G(x, s)$ by .02 by taking a step in the
    primal variables in the direction $-d(e^{T}\bar{s})$.

    b) If $\|d(e^{T}\bar{s})\|_2 < .22$, we can decrease $G(x, s)$ by .02 by replacing
    $\bar{s}$ by $\tilde{s} = s(\tilde{\Delta})$ where $\tilde{\Delta} \in (0, e^{T}\bar{s})$ satisfies $\|d(\tilde{\Delta})\|_2 = .22$.

This leads to the following algorithm for solving LP. Let (A, b, c) be the data for the LP, let $\bar{x}$ and $(\bar{\pi}, \bar{s})$ be initial primal and dual solutions, where $\bar{x} > 0$ and let $\mu > 0$ be the optimality tolerance. We have:

<u>Algorithm 2.</u> $(A, b, c, \bar{x}, \bar{s}, \bar{\pi}, \mu)$

<u>Step 0. Initialize.</u>

    Set $q = n + \sqrt{n}$, $\gamma = .22$

<u>Step 1. Test for Optimality</u>

    If $\bar{x}^{T}\bar{s} \leq \mu$, stop.

<u>Step 2. Rescale</u>

    Set $\bar{A} = A\bar{X}$, $\bar{c} = \bar{X}c$, $\bar{t} = \bar{X}\tilde{s}$.

Step 3. Compute direction function

$$d(\Delta) \;=\; [I - \overline{A}^{T}(\overline{A}\,\overline{A}^{T})^{-1}\overline{A}]\left[\left(\frac{q}{\Delta}\right)\overline{t} - e\right]$$

If $\|d(e^{T}\overline{t})\|_2 \;\geq\; \gamma,$ go to Step 5. Otherwise go to Step 4.

Step 4. Recompute Dual Variables

Solve for $\widetilde{\Delta} \in (0, e^{T}\overline{t})$ such that $\widetilde{\Delta}$ is the smallest value of Δ for which $\|d(\widetilde{\Delta})\|_2 = \gamma.$

Set $\widetilde{t} = \left(\dfrac{\widetilde{\Delta}}{q}\right)\!\left(e + d(\widetilde{\Delta})\right)$ and $\widetilde{\pi} = (\overline{A}\,\overline{A}^{T})^{-1}\overline{A}\left[\overline{c} - \left(\dfrac{\widetilde{\Delta}}{q}\right)e\right]$

Set $\overline{\pi} = \widetilde{\pi}$ and $\overline{s} = \overline{X}^{-1}\widetilde{t}$, and go to Step 1.

Step 5. Take step in Primal variable space.

Set $\overline{d} = d(e^{T}\overline{t})$. Set $\alpha = 1 + \gamma - \dfrac{1}{\sqrt{1 + 2\lambda}}$.

Set $\widetilde{z} = e - \alpha\,\overline{d}/\|\overline{d}\|_2.$

Set $\overline{x} = \overline{X}\widetilde{z}$, and return to Step 1.

Note that in Step 5, $\alpha = (1/6)$ because $\gamma = .22$.

Regarding the complexity of Algorithm 2, we have:

Theorem 4.1 (Complexity of Algorithm 2). If (x^0, s^0) are the initial values of $\overline{x}$ and $\overline{s}$ in Algorithm 2 and $G(x^0, s^0) \leq O(\sqrt{n}\,L)$, then Algorithm 2 solves LP in $O(\sqrt{n}\,L)$ iterations.

Proof: The proof is an immediate consequence of Proposition 2.1. From Remark 4.1 and the invariance of the potential function $G(x, s)$ under scaling, we have that at each iteration $G(x, s)$ is reduced by at least $\delta = .02.$ Because $q = n + \sqrt{n}$, then $(q - n) = \sqrt{n}$, and so by Proposition 2.1, Algorithm 2 converges in $O(\sqrt{n}\,L)$ iterations. ∎

Note Algorithm 2 is almost the same as Algorithm 1, but with a different value of γ in Step 0 and consequently a different step length α in Step 5. The smaller value of γ in Algorithm 2 suggests that it will update dual variables less frequently than in Algorithm 1, and so will control the influence of the duality gap $\Delta = x^T s$ on the projected gradient direction $d(\Delta)$ less exactly. Perhaps this is a contributing factor to the improved complexity bound.

Most of the remarks that follow Algorithm 1 at the end of Section 3 are also valid for Algorithm 2. If the algorithm recomputes dual variables in Step 4, one can then proceed directly to Step 5 because $d(e^T s) \geq \gamma$ at the end of Step 4 and there is no need to recompute the least-squares solution of Step 3. As in Algorithm 1, one can augment Steps 4 and 5 by a line-search of the potential function $G(x, s)$. Also, one need not start the algorithm with a dual feasible solution. In fact, only a lower bound on the optimal objective value is needed (see Section 3 for details).

The algorithm of Ye [18] was the first (and at this point, the only other) algorithm with a complexity bound of $O(\sqrt{n}\, L)$ iterations that does not explicitly require a condition that the iterates lie near the central trajectory. His algorithm is similar to Algorithm 2 in many ways. First, both algorithms use primal scaling only, compute a direction function based on the projected gradient of the potential function, and take either a step in the primal space or a step in the dual slack space. However, the control sequence for when to take a primal versus a dual step is different in the two algorithms. Even if both algorithms utilize a line-search, they still may perform differently. This is because of the control logic for when to take a primal or a dual step.

5. On the Optimality of $q = n + \sqrt{n}$ in Algorithm 2

In Section 4, we have presented an $O(\sqrt{n}\, L)$ iteration algorithm for LP, by choosing the value $q = n + \sqrt{n}$ in the primal-dual potential function $G(x, s)$ of (2.2). In this section, we consider modifying Algorithm 2 in two ways: by setting $q = n + n^t$, where $t \geq 0$, and by setting $\gamma = .22n^{-k}$ where $k \geq 0$. We will prove:

Theorem 5.1 (Complexity of Algorithm 2 with $q = n + n^t$)

If $q = n + n^t$ is used in Algorithm 2 and γ is chosen by:

$$\gamma = \begin{cases} .22 & \text{if } t \geq 1/2 \\ .22n^{t-1/2} & \text{if } t \leq 1/2 \ , \end{cases}$$

then the complexity of the algorithm if $O(n^r \ L)$ iterations, where $r = \max\{t, 1-t\}$.

Remark 5.1. The value of t that minimizes the complexity bound of the number of iterations is $t = 1/2$, and so Algorithm 2 with $q = n + \sqrt{n}$ and $\gamma = .22$ is optimal in this measure. A different discussion of the suggested optimality of $q = n + \sqrt{n}$ (in a different context) can be found in Todd and Ye [14].

Remark 5.2. Algorithm 2 can be executed with $q = n + 1$, with a complexity of bound $O(n \ L)$ iterations, by setting $t = 0$ and $\gamma = .22/\sqrt{n}$. This improves the complexity bound of Gonzaga [7], whose results imply an $O(n^2 \ L)$ iteration bound if $q = n + 1$ for a primal potential function as in (2.1).

We will first prove Theorem 5.1 for the case $t \geq 1/2$. The case $t \leq 1/2$ is a bit more involved.

Proof of Theorem 5.1 for $t \geq 1/2$:

Suppose Algorithm 2 is implemented with $q = n + n^t$ and $\gamma = .22$. Then if the algorithm takes a primal step (Step 5), Proposition 4.1 ensures that the potential function $G(x, s)$ decreases by at least $\delta = .02$, as this result is independent of the value of q.

Suppose instead that Algorithm 2 takes a dual step (Step 4). By rescaling if necessary, we assume that the current primal and dual slack vectors are $(\bar{x}, \bar{s}) = (e, \bar{s})$ and let $\tilde{s}$ be

the recomputed dual slack vector. Retracing the proof of Lemma 4.4, we reach line (4.7) which states:

$$G(e,\tilde{s}) - G(e,\bar{s}) \le - (q - n)\ \frac{(q - n - .22\sqrt{n})}{q} + .367 \qquad (5.1)$$

$$= \frac{n^t (n^t - .22\sqrt{n})}{n + n^t} + .367$$

$$\le - \frac{.78n^{2t}}{n + n^t} + .367 \ . \qquad (5.2)$$

If $t \le 1$, the denominator above is bounded from above by $2n$, so $G(e,\tilde{s}) - G(e,\bar{s}) \le - .39n^{2t-1} + .367 \le - .39 + .367 \le - (.02)$. If $t \ge 1$, the denominator above is bounded from above by $2n^t$, so $G(e,\tilde{s}) - G(e,\bar{s}) \le - .39n^t + .367 \le - (.02)$ if $t \ge 1$. In either case, we obtain a dual improvement of at least $\delta = .02$. Thus the overall complexity of the algorithm, from Proposition 2.1, is $O(n^t L)$. ∎

We now proceed to piece together the proof Theorem 5.1 in the case $t \le 1/2$. By rescaling if necessary, we assume that the current primal and dual slack vectors are $(\bar{x},\bar{s}) = (e,\bar{s})$. We first analyze a primal step.

Proposition 5.1 (Primal Step Improvement) Let $\bar{d} = d(e^{\bar{T}s})$. Suppose

$\|d\|_2 \ge \gamma = .22n^{t - 1/2}$. Then by setting $\alpha = 1 - \dfrac{1}{\sqrt{1 + 2\gamma}}$,

we obtain

$$G(e - \alpha\bar{d}/\|\bar{d}\|_2, \bar{s}) \le G(e,\bar{s}) - .02n^{2t - 1} \ .$$

Proof. The proof proceeds identically to that of Proposition 3.1, which is of the same form (with $.8$ replaced $\gamma = .22n^{t-1/2}$) . Tracing the proof of Proposition to line (3.2) we obtain.

$$G(e - \alpha\bar{d}/\|\bar{d}\|_2, \bar{s}) - G(e,\bar{s}) \le - \alpha\|\bar{d}\|_2 + \frac{\alpha^2}{2(1 - \alpha)} \ .$$

However, $\|\bar{d}\|_2 \geq \gamma = .22n^{t-1/2}$. Substituting this value of γ in Proposition A.2 of the Appendix, we obtain

$$G(e - \alpha\bar{d}/\|\bar{d}\|_2 , \bar{s}) - G(e,\bar{s}) \leq -.02n^{2t-1} \quad \blacksquare$$

We now analyze a dual step improvement. We proceed by modifying Lemmas 4.2 and 4.3.

Lemma 5.2. If $s = \dfrac{\Delta}{q}(e + d)$, where $\|d\|_2 < \gamma < 1$, then

$$\| s - \left(\frac{e^Ts}{e}\right)e \|_2 < \left(\frac{2\gamma}{1-\gamma}\right)\left(\frac{e^Ts}{n}\right) . \quad \blacksquare$$

Lemma 5.3. If $s = \dfrac{\Delta}{q}(e + d)$, where $\|d\|_2 < \gamma < 1/3$, then

$$\sum_{j=1}^{n} \ln(s_j) \geq n\ln\left(\frac{e^Ts}{n}\right) - \frac{\beta^2}{2(1-\beta)} , \quad \text{where } \beta = \frac{2\gamma}{1-\gamma} \quad \blacksquare$$

The proofs of Lemmas 5.2 and 5.3 are identical to the proofs of Lemmas 4.2 and 4.3, with γ substituted for .22 in Lemmas 4.2 and 4.3, and β substituted for .565 in Lemma 4.3.

We now are in a position to prove a result regarding potential function improvement if we take a dual step.

Lemma 5.4 (Dual Step Improvement). Suppose $q = n + n^t$ where $0 \leq t \leq 1/2$. Suppose $\|d(e^{T}\bar{s})\|_2 < \gamma = .22n^{t-1/2}$, and that c^Tx is not constant on the feasible region of LP. Let $\widetilde{\Delta} \in (0, e^{T}\bar{s})$ satisfy $\|d(\widetilde{\Delta})\|_2 = \gamma$ and let $\widetilde{\pi} = \pi(\widetilde{\Delta})$ and $\widetilde{s} = s(\widetilde{\Delta})$. Then $(\widetilde{\pi}, \widetilde{s})$ are feasible for LD, and

$$G(e,\widetilde{s}) - G(e,\bar{s}) \leq -.02n^{2t-1} .$$

Proof: Parallelling the proof of Lemma 4.4, we obtain

$$G(e, \tilde{s}) - G(e, \bar{s}) = q \ln\left(\frac{e^{T\tilde{s}}}{e^{T\bar{s}}}\right) - \sum_{j=1}^{n} \ln(\tilde{s}_j) + \sum_{j=1}^{n} \ln(\bar{s}_j) \ . \tag{5.3}$$

Combining the above with inequality (4.5) and Lemma 5.3, we obtain

$$G(e, \tilde{s}) - G(e, \bar{s}) \leq q \ln\left(\frac{e^{T\tilde{s}}}{e^{T\bar{s}}}\right) - n \ln\left(\frac{e^{T\tilde{s}}}{n}\right) + \frac{\beta^2}{2(1-\beta)} + n \ln\left(\frac{e^{T\tilde{s}}}{n}\right)$$

$$= (q - n) \ln\left(\frac{e^{T\tilde{s}}}{e^{T\bar{s}}}\right) + \frac{\beta^2}{2(1-\beta)} \ , \tag{5.4}$$

where $\beta = \frac{2\gamma}{1-\gamma}$. However, $\frac{\beta^2}{2(1-\beta)} = \frac{2\gamma^2}{(1-\gamma)(1-3\gamma)}$ and $\gamma = .22n^{t-1/2} \leq .22$, so that

$$\frac{\beta^2}{2(1-\beta)} \leq \frac{2\gamma^2}{(.78)(1-.66)} \leq .365n^{2t-1} . \tag{5.5}$$

Next, notice that exactly as in the proof Lemma 4.4, we have that

$$e^{T\tilde{s}} = \left(\frac{\tilde{\Delta}}{q}\right)\left(n + e^{T}d(\tilde{\Delta})\right) \leq \left(\frac{e^{T\bar{s}}}{q}\right)\left(n + \sqrt{n}\|d(\tilde{\Delta})\|_2\right) = \left(\frac{e^{T\bar{s}}}{q}\right)\left(n + .22n^{t-1/2+1/2}\right)$$

and hence $\left(\frac{e^{T\tilde{s}}}{e^{T\bar{s}}}\right) \leq \frac{n + .22n^t}{q} = 1 - \frac{n^t - .22n^t}{q}$,

whereby $(q - n) \ln\left(\frac{e^{T\tilde{s}}}{e^{T\bar{s}}}\right) \leq \frac{n^t(n^t - .22n^t)}{q}$

$$\leq \frac{.78n^{2t}}{2n} = .39n^{2t-1} . \tag{5.6}$$

Combining (5.4), (5.5) and (5.6) yields

$$G(e, \tilde{s}) - G(e, \bar{s}) \leq -.39n^{2t-1} + .365n^{2t-1} \leq -(.02)n^{2t-1}. \blacksquare$$

Proof of Theorem 5.1 for $t \leq 1/2$: From Proposition 5.1 and Lemma 5.4, we have that we can achieve a decrease in $G(x, s)$ of at least $\delta = .02n^{2t-1}$. Thus, the overall complexity of the algorithm, from Proposition 2.1, is $O(n^t n^{1-2t} L) = O(n^{1-t} L). \blacksquare$

Remark 5.3. The choice of γ in Theorem 5.1 obviously influences the complexity bound. One can easily verify that the potential function improvement in a primal step is $O(\gamma^2)$. Thus, a large value of γ is desirable. However, if $\gamma > O(n^{t-1/2})$ and $t < 1/2$, then a potential function improvement in a dual step cannot be guaranteed. The formula for γ in Theorem 5.1 is the largest (and best) value of γ with respect to minimizing the complexity bound of the algorithm.

6. Concluding Remarks

Algorithm Complexity. The overall complexity of Algorithm 2 is $O(n^{3.5} L)$ operations, because at each of the $\sqrt{n} L$ iterations an m x m system of equations must be solved, which requires $O(n^3)$ operations. However, by solving the system inexactly using, for example, the updating algorithm of Gonzaga [6] as presented in [7], it should be possible to reduce the overall complexity by a factor of $\sqrt{n}$ to $O(n^3 L)$ iterations. However, this modification will probably not be very useful in practice, as it limits the use of a line-search of the potential function.

Problems in Arbitrary Form. Algorithms 1 and 2 are stated for LP problems in standard form. For problems in other forms, the implementation of the algorithm is still straightforward. Suppose that we wish to solve:

$$\text{minimize} \quad c^T x$$

$$\begin{aligned}
\text{s.t.} \quad & Ax = b \\
& Gx - v = h \\
& v \geq 0 \quad .
\end{aligned}$$

This is a quite general form and encompasses problems with upper and lower bounds, etc., by creating G, h with appropriate submatrices of the identify matrix, etc. The primal-dual potential function then is

$$H(v, s) = \ln(v^T s) - \sum_{j=1}^{n} \ln(v_j) - \sum_{j=1}^{n} \ln (s_j)$$

where (π, s) must satisfy dual interior feasibility, namely

$$A^T \pi + G^T s = c$$
$$s > 0 .$$

Scaling is done with the primal slack variables. If $(\bar{x}, \bar{v})$ are primal feasible, the LP is rescaled to

$$\text{minimize} \quad c^T x$$

$$\text{s.t.} \quad Ax \qquad = b$$
$$\bar{V}^{-1} Gx - t = \bar{V}^{-1} h$$
$$t \geq 0$$

and the projection of the gradient of $H(v, s)$ in the rescaled space is then computed.

Appendix

Proposition A.1. (see Karmarkar [8] and Todd and Ye [14]).

(i) $\ln(1+x) \leq x$ for $x > -1$.

(ii) If $|x| \leq \alpha < 1$, then $\ln(1+x) \geq x - \dfrac{x^2}{2(1-\alpha)}$.

Proof. (i) follows from the concavity of the log function.

For (ii) note that if $|x| < 1$, $\ln(1+x) = x - \dfrac{x^2}{2} + \dfrac{x^3}{3} - \dfrac{x^4}{4} + \cdots$

$$\geq x - \frac{x^2}{2} - \frac{|x|^3}{2} - \frac{|x|^4}{2} - \cdots$$

$$= x - \frac{x^2}{2(1-|x|)} \geq x - \frac{x^2}{2(1-\alpha)}. \qquad \blacksquare$$

Proposition A.2 (See Freund [4], Proposition 4.1) If

$$f(\alpha) = -\gamma\alpha + \frac{\alpha^2}{2(1-\alpha)},$$

where $\gamma \leq .22$, then

with $\overline{\alpha} = 1 - \dfrac{1}{\sqrt{1+2\gamma}}$, $f(\overline{\alpha}) \leq -\dfrac{.02\gamma^2}{(.22)^2}$.

Proof: Direct substitution shows $f(\overline{\alpha}) = -1 - \gamma + \sqrt{1+2\gamma}$. However, the function

$\dfrac{1+\gamma-\sqrt{1+2\gamma}}{\gamma^2}$ is decreasing in γ. Therefore, if $\gamma \leq .22$,

$$\frac{f(\overline{\alpha})}{\gamma^2} = -\frac{1+\gamma-\sqrt{1+2\gamma}}{\gamma^2} \leq -\frac{(1.22-\sqrt{1+2(.22)})}{(.22)^2} = -\frac{.02}{(.22)^2}. \qquad \blacksquare$$

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] Bayer, D. A., and J.C. Lagarias. 1987. The nonlinear geometry of linear programming, I. Affine and projective scaling trajectories, II. Legendre transform coordinates and central trajectories. **Transactions of the American Mathematical Society,** forthcoming.

[4] Freund, R. 1988. Projective transformations for interior point methods, part I: Basic theory and linear programming. M.I.T. Operations Research Center working paper OR 179-88.

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

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

[7] Gonzaga, C.C. 1988. Polynomial affine algorithms for linear programming. Report ES-139/88, Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brazil.

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

[9] Megiddo, N. 1986. Pathways to the optimal set in linear programming. Research Report RJ 5295, IBM Almaden Research Center, San Jose, Ca.

[10] 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.

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

[12] 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.

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

[14] Todd, M.J., and Y. Ye. 1987. A centered projective algorithm for linear programming. Technical Report 763, School of ORIE, Cornell University, Ithaca, New York.

[15] Vaidya, P. 1987. An algorithm for linear programming which requires $O(((m + n)n^2 + (m + n)^{1.5} n)L)$ arithmetic operations. AT&T Bell Laboratories, Murray Hill, N.J.

[16] Ye, Yinyu. 1987. Interior algorithms for linear, quadratic, and linearly constrained convex programming. Ph. D. thesis, Department of Engineering Economic Systems, Stanford University, Stanford, Ca.

[17] Ye, Yinyu. 1988. A class of potential functions for linear programming. Presented at the Joint Summer Research Conference: Mathematical developments arising from linear programming algorithms, Bowdoin College, Brunswick, Me.

[18] Ye, Yinyu. 1988. A class of potential functions for linear programming. Department of Management Sciences, The University of Iowa, Iowa City, Iowa.