An interior point method, based on rank-one updates, for linear programming

Jos F. Sturm *
Shuzhong Zhang †

Revised on December 1, 1995

Abstract

We propose a polynomial time primal-dual potential reduction algorithm for linear programming. Unlike any other interior point method, the new algorithm is based on a rank-one updating scheme for sequentially computing the projection matrices. For a standard linear programming problem, the number of operations required is $\mathcal{O}(mn)$ per main iteration and the overall computational complexity is $\mathcal{O}(mn^{2.5}L)$.

Key words: linear programming, interior point method, potential function.

1. Introduction

In his seminal paper [1], Karmarkar proposed a projective potential reduction method that solves linear programs in $\mathcal{O}(nL)$ main iterations, where n stands for the number of inequality constraints, and L is the input length of the problem. A single iteration of Karmarkar's algorithm requires solving a system of linear equations, which takes $\mathcal{O}(n^3)$ operations in the worst case. By using incomplete updates of this equation system, the average work per step can be reduced to $\mathcal{O}(n^{2.5})$. Hence, Karmarkar's algorithm achieves an overall complexity of $\mathcal{O}(n^{3.5}L)$ for linear programming.

In this paper, we propose a potential reduction algorithm using the Tanabe-Todd-Ye primal-dual potential function [2, 3]. The new algorithm has the feature that at each step the matrix of the afore mentioned system of linear equations differs from the previous matrix by only a rank-one update. As a result, the work per iteration is reduced to $\mathcal{O}(n^2)$ operations. However, a rank-one update is in general less effective in reducing the potential function value than a full rank update. To be more precise, we arrive at an $\mathcal{O}(n^{1.5}L)$ worst case iteration bound. Hence, the new algorithm needs $\mathcal{O}(n^{3.5}L)$ operations in total, just like Karmarkar's algorithm. At each iteration of the new method however, the amount of required operations is comparable to that of one simplex pivot step.

The paper is organized as follows. In Section 2, we briefly discuss the primal-dual geometry of linear programming. The effect of a rank-1 update on the primal-dual solution pair is derived in Section 3. Subsequently, the new potential reduction algorithm is stated in

* Tinbergen Institute Rotterdam, The Netherlands, sturm@opres.few.eur.nl.
† Erasmus University Rotterdam, The Netherlands, zhang@opres.few.eur.nl.

Section 4, and we proceed in Section 5 by deriving the complexity result. We conclude the paper in Section 6.

2. Problem and discussions

In this paper we consider the standard linear program:

$$(P) \quad \min_x \{c^{\mathrm{T}}x \mid Ax = b, x \geq 0\}$$

and its dual

$$(D) \quad \max_{y,s} \{b^{\mathrm{T}}y \mid A^{\mathrm{T}}y + s = c, s \geq 0\}$$

where A is an $m \times n$ matrix, $n \geq 2$, and b and c are vectors of dimension m and n respectively.

We make the standard assumptions that A has full row rank and that (P) and (D) have interior feasible solutions.

For given positive vector $d \in \Re^n_{++}$, let

$$\mathcal{F}_P(d) := \{x \mid ADx = b\}$$

where

$$D := \mathrm{diag}(d_1, d_2, \ldots, d_n)$$

and let

$$\mathcal{F}_D(d) := \{s \mid DA^{\mathrm{T}}y + s = Dc \text{ for some } y \in \Re^m\}.$$

Denoting the all-one vector by e, it is easily seen that primal feasibility amounts to

$$x \in \mathcal{F}_P(e) \text{ and } x \geq 0$$

and dual feasibility is equivalent to

$$s \in \mathcal{F}_D(e) \text{ and } s \geq 0.$$

As $\mathcal{F}_P(d)$ and $\mathcal{F}_D(d)$ are orthogonal complementary affine spaces, they share a single intersection point, say v_d:

$$\{v_d\} = \mathcal{F}_P(d) \cap \mathcal{F}_D(d).$$

Let

$$x_d := Dv_d$$

and

$$s_d := D^{-1}v_d.$$

Obviously, $x_d \in \mathcal{F}_P(e)$ and $s_d \in \mathcal{F}_D(e)$. Because of the positivity of d, we also have

$$x_d \geq 0 \text{ and } s_d \geq 0$$

if and only if

$$v_d \geq 0.$$

A natural question arises: Is v_d always nonnegative for positive d? Or equivalently: Are x_d and s_d always feasible? As expected, the answer is negative. A simple example is:

Example 2.1. *Let* $A = \begin{bmatrix} 1 & 2 \end{bmatrix}, b = \begin{bmatrix} 4 \end{bmatrix}$ *and* $c^T = \begin{bmatrix} 0 & 7 \end{bmatrix}$. *Then, choosing* $d^T = \begin{bmatrix} 1 & 1 \end{bmatrix}$ *yields*

$$v_d = \begin{bmatrix} -2 \\ 3 \end{bmatrix},$$

which has a negative component.

If, on the other hand, interior feasible solutions $x \in \mathcal{F}_P(e)$, $x > 0$ and $s \in \mathcal{F}_D(e)$, $s > 0$ are known, then it is easy to construct d such that $v_d > 0$. Namely, letting

$$d_i = \sqrt{x_i/s_i} \text{ for } i = 1, 2, \ldots, n$$

yields

$$(v_d)_i = \sqrt{x_i s_i} \text{ for } i = 1, 2, \ldots, n.$$

This triggers the next question: Assuming that we have a pair of interior feasible solutions, is it possible to update the related vector d such that the feasibility is retained and, at the same time, the solutions are improved?

The goal of this paper is to give an assertive answer to the above question. Namely, we propose an algorithm that generates a sequence $\{d^k \mid k = 0, 1, 2, \cdots\}$ for which the associated sequence of primal-dual feasible pairs (x_{d^k}, s_{d^k}) converges to a pair of optimal solutions for (P) and (D).

First we derive formulas to express v_d, x_d and s_d for a given vector $d \in \Re^n_{++}$. Let Q_d and P_d denote the projection matrices onto the image of DA^{T} and the kernel of AD respectively, i.e.

$$Q_d := DA^{\mathrm{T}}(AD^2A^{\mathrm{T}})^{-1}AD$$

and

$$P_d := I - Q_d.$$

By definition, $v_d \in \mathcal{F}_P(d) \cap \mathcal{F}_D(d)$, i.e.

$$ADv_d = b \tag{1}$$

and

$$DA^{\mathrm{T}}y_d + v_d = Dc, \tag{2}$$

for some $y_d \in \Re^m$. Pre-multiplying (1) with $DA^{\mathrm{T}}(AD^2A^{\mathrm{T}})^{-1}$ yields

$$Q_d v_d = DA^{\mathrm{T}}(AD^2A^{\mathrm{T}})^{-1}b$$

and pre-multiplying (2) with P_d yields

$$P_d v_d = P_d Dc.$$

Hence,

$$v_d = Q_d v_d + P_d v_d = DA^{\mathrm{T}}(AD^2A^{\mathrm{T}})^{-1}b + P_d Dc,$$

which satisfies (1) and (2) with

$$y_d = (AD^2A^{\mathrm{T}})^{-1}(AD^2c - b). \tag{3}$$

3. Single component update

In this section, we analyze the effect of a change in a single component of d on the associated pair (x_d, s_d). Suppose that we update component d_j, for some $j \in \{1, 2, \cdots, n\}$. For a step size $t > -1$, let the new vector of scales $d(t)$ be defined as follows:

$$d_j(t) = \sqrt{1+t}\,d_j$$

and

$$d_i(t) = d_i \text{ for } i \in \{1, 2, \cdots, n\} \backslash \{j\},$$

i.e. the components other than j remain unchanged.

We define

$$D(t) := \operatorname{diag}(d_1(t), d_2(t), \ldots, d_n(t)).$$

Let

$$u_d := (AD^2A^{\mathrm{T}})^{-1}ADe_j,$$

where e_j is the j-th unit vector. Using the rank-1 updating formula of Sherman, Morrison and Woodbury, we obtain

$$(AD(t)^2A^{\mathrm{T}})^{-1} = (AD^2A^{\mathrm{T}})^{-1} - \frac{t}{1 + t\,\|Q_d e_j\|^2} u_d u_d^{\mathrm{T}}.$$

Furthermore,

$$AD(t)^2c = AD^2c + t(d_jc_j)ADe_j.$$

It follows using (3) that

$$\begin{aligned} y_{d(t)} &= (AD(t)^2A^{\mathrm{T}})^{-1}(AD(t)^2c - b) \\ &= y_d + t(d_jc_j)u_d \\ &\quad -t\frac{t(d_jc_j)\|Q_de_j\|^2 + u_d^{\mathrm{T}}(AD^2c-b)}{1+t\|Q_de_j\|^2}u_d. \end{aligned}$$

From (3) and (2) we have

$$u_d^{\mathrm{T}}(AD^2c - b) = e_j^{\mathrm{T}}DA^{\mathrm{T}}y_d = d_jc_j - (v_d)_j,$$

and therefore

$$y_{d(t)} = y_d + \frac{t(v_d)_j}{1+t\|Q_de_j\|^2}u_j.$$

Now it follows that

$$\begin{aligned} Ds_{d(t)} &= Dc - DA^{\mathrm{T}}y_{d(t)} \\ &= v_d - \frac{t}{1+t\|Q_de_j\|^2}(v_d)_jQ_de_j. \end{aligned} \tag{4}$$

As $x_{d(t)} = D(t)v(t) = D(t)^2s_{d(t)}$, we have

$$\begin{aligned} D^{-1}x_{d(t)} &= (D^{-2}D(t)^2)Ds_{d(t)} \\ &= (I + te_je_j^{\mathrm{T}})Ds_{d(t)} \\ &= v_d - \frac{t}{1+t\|Q_de_j\|^2}(v_d)_jQ_de_j \\ &\quad +t(v_d)_je_j - \frac{t^2\|Q_de_j\|^2}{1+t\|Q_de_j\|^2}(v_d)_je_j \\ &= v_d + \frac{t}{1+t\|Q_de_j\|^2}(v_d)_jP_de_j. \end{aligned} \tag{5}$$

We introduce $\alpha(t)$,

$$\alpha(t) := \frac{t}{1+t\|Q_de_j\|^2},$$

which is a strictly increasing function for $t > -1$. This function maps the interval $(-1, \infty)$ into the interval

$$\{\alpha \mid \|P_de_j\|^2\alpha > -1 \text{ and } \|Q_de_j\|^2\alpha < 1\}.$$

The inverse transformation of $\alpha(t)$ is

$$t(\alpha) = \frac{\alpha}{1 - \alpha \|Q_d e_j\|^2}. \tag{6}$$

From now on, we consider α as the step size parameter, and we compute $t(\alpha)$ from (6). As a matter of notation, we write

$$x(\alpha) := x_{d(t(\alpha))}, s(\alpha) := s_{d(t(\alpha))}$$

and

$$v(\alpha) := v_{d(t(\alpha))}.$$

If no confusion is possible, we will dispense with the argument d and simply write P, Q, v, x, s and y to denote P_d, Q_d, v_d, x_d, s_d and y_d. Now we rewrite (5) and (4) as follows:

$$D^{-1}x(\alpha) = v + \alpha v_j P e_j \tag{7}$$

and

$$Ds(\alpha) = v - \alpha v_j Q e_j. \tag{8}$$

These relations imply that

$$\begin{aligned} \|v(\alpha)\|^2 &= x(\alpha)^{\mathrm{T}} s(\alpha) \\ &= (v + \alpha v_j P e_j)^{\mathrm{T}} (v - \alpha v_j Q e_j) \\ &= \|v\|^2 + \alpha v_j e_j^{\mathrm{T}} (P - Q) v. \end{aligned} \tag{9}$$

Hence, the duality gap between $x(\alpha)$ and $s(\alpha)$ is linear in α.

As is common in the interior point methodology, we restrict ourselves to positive, i.e. interior, solutions $x > 0$ and $s > 0$. Introducing the quantity

$$v_{\min} := \min\{v_1, v_2, \cdots, v_n\},$$

we obtain from (7) and (8) that

$$x(\alpha) > 0 \text{ and } s(\alpha) > 0$$

as long as $\mid \alpha \mid < v_{\min}/v_j$.

4. A potential reduction algorithm

Consider for $x > 0$ and $s > 0$ the Tanabe-Todd-Ye primal-dual potential function

$$\phi(x,s) = (n+l)\log(x^{\mathrm{T}}s) - \sum_{i=1}^{n}\log(x_i s_i) - n\log\sqrt{n},$$

where we let $l := \sqrt{2n}$ in this paper.

From the arithmetic-geometric mean inequality, it is easily seen that

$$\phi(x,s) \geq \sqrt{2n}\log(x^{\mathrm{T}}s),$$

and equality holds only if $v = X^{1/2}S^{1/2}e$ is a positive multiple of the all-one vector e. This observation is the essence of the proof of the following lemma, which is due to Ye [3].

Lemma 4.1. *If an algorithm reduces $\phi(x,s)$ by at least a fixed quantity $\delta > 0$ in each iteration, then the algorithm can be used to solve the pair (P), (D) in $\mathcal{O}(\sqrt{n}L/\delta)$ iterations.*

Let

$$p_{ij} := e_i^{\mathrm{T}}Pe_j \text{ for } i = 1,2,\ldots,n$$

and

$$q_{ij} := e_i^{\mathrm{T}}Qe_j \text{ for } i = 1,2,\ldots,n.$$

For $\mid \alpha \mid < v_{\min}/v_j$, we have from (7), (8) and (9) that

$$\begin{aligned}\phi(x(\alpha),s(\alpha)) &= \phi(x,s) + (n+\sqrt{2n})\log(1+\alpha\frac{v_j e_j^{\mathrm{T}}(P-Q)v}{\|v\|^2}) \\ &\quad -\sum_{i=1}^{n}\log(1+\frac{\alpha v_j}{v_i}p_{ij}) - \sum_{i=1}^{n}\log(1-\frac{\alpha v_j}{v_i}q_{ij}).\end{aligned} \tag{10}$$

Let

$$r := (P-Q)(\frac{(n+\sqrt{2n})}{\|v\|^2}v - V^{-1}e), \tag{11}$$

where

$$V := \operatorname{diag}(v_1, v_2, \ldots, v_n).$$

It is easily seen from (10) that

$$\phi(x(\alpha),s(\alpha)) = \phi(x,s) + \alpha v_j r_j + o(\alpha).$$

We propose to select the index j such that $\mid r_j \mid = \|r\|_\infty$. This choice results in the following algorithm.

Algorithm 1. *(A,b,c,d^0)*

The initial scaling $d^0 > 0$ is chosen such that $v_{d^0} > 0$.

Step 0 Initialization. *Set $k = 0$.*

Step 1 Optimality test. *Let $d = d^k$. Stop if based on (x_d, s_d) an optimal pair (x^*, s^*) can be found.*

Step 2 Choose the updating index. *Compute r according to (11). Set*

$$j = \arg\max_{1 \le i \le n} \mid r_i \mid$$

Step 3 Step size choice. *Compute α such that*

$$\phi(x_d(\alpha), s_d(\alpha)) < \phi(x_d, s_d) - \frac{1}{6(n+2)}$$

Step 4 Take step. *Set $d^{k+1} = d + (\sqrt{1 + t(\alpha)} - 1) d_j e_j$.*

Step 5 *Set $k = k + 1$ and return to Step 1.*

Notice that in order to perform the algorithm it is necessary to iteratively compute v and r. To do this, it is sufficient to update the matrices $(AD^2A^{\mathrm{T}})^{-1}$ and $(AD^2A^{\mathrm{T}})^{-1}AD$. Using the rank-one updating formula, this will take $\mathcal{O}(mn)$ operations, whereas a main iteration of other interior point algorithms typically requires $\mathcal{O}(m^2n)$ operations in the worst case. Also notice that the index selection rule, i.e. Step 2, is similar to the most negative pivot rule of the simplex method.

5. Convergence analysis

In estimating the reduction $\phi(x, s) - \phi(x(\alpha), s(\alpha))$, we shall make use of the inequality

$$\rho \ge \log(1 + \rho) \ge \rho - \frac{\rho^2}{2(1- \mid \rho \mid)} \text{ for any } \rho \in (-1, 1), \tag{12}$$

which is proven by Karmarkar [1].

Applying (12) to (10) yields for $\mid \alpha \mid < v_{\min}/v_j$

$$\begin{aligned} \phi(x, s) - \phi(x(\alpha), s(\alpha)) &\ge -\alpha v_j r_j \\ &\quad -\alpha^2 \frac{v_j^2}{v_{\min}^2} \frac{\|Pe_j\|^2 + \|Qe_j\|^2}{2(1- \mid \alpha \mid v_j/v_{\min})} \\ &= -\alpha v_j r_j - \frac{(\alpha v_j/v_{\min})^2}{2(1- \mid \alpha \mid v_j/v_{\min})}. \end{aligned} \tag{13}$$

Due to the index selection rule in Algorithm 1, there holds $\mid r_j \mid = \|r\|_\infty$. The following lemma provides a lower bound on $\|r\|_\infty$.

Lemma 5.1. *It holds that*

$$\|r\|_\infty > \frac{1}{\sqrt{n} v_{\min}}.$$

Proof:

Obviously, we have for any $w \in \Re^n$ that

$$\|(P-Q)w\| = \|(P+Q)w\| = \|w\|.$$

Therefore, we have

$$\|r\|^2 = \left\| \frac{(n+\sqrt{2n})}{\|v\|^2} v - V^{-1}e \right\|^2 = \left\| \frac{\sqrt{2n}}{\|v\|^2} v + \frac{n}{\|v\|^2} v - V^{-1}e \right\|^{2}.$$

Using $v \perp \frac{n}{\|v\|^2} v - V^{-1}e$, it follows that

$$\begin{aligned}
\|r\|^2 &= \frac{2n}{\|v\|^2} + \left\| \frac{n}{\|v\|^2} v - V^{-1}e \right\|^2 \\
&\geq \frac{2n}{\|v\|^2} + \left(\frac{n}{\|v\|^2} v_{\min} - \frac{1}{v_{\min}}\right)^2 \\
&= \frac{1}{v_{\min}^2} + \frac{n^2 v_{\min}^2}{\|v\|^4}. \\
&> \frac{1}{v_{\min}^2}. \qquad (14)
\end{aligned}$$

Since $\|r\|_\infty \geq \|r\| / \sqrt{n}$, (14) implies the desired result.

□

Combining Lemma 5.1 and (13) we obtain a lower bound on the reduction $\phi(x,s) - \phi(x(\alpha), s(\alpha))$:

Lemma 5.2. *For* $\alpha = -\frac{|r_j|}{r_j} \frac{2v_{\min}/v_j}{3\sqrt{n+2}}$, *we have*

$$\phi(x,s) - \phi(x(\alpha), s(\alpha)) > \frac{1}{3(n+2)}.$$

Proof:

From (13) and the fact that $| r_j | = \|r\|_\infty$ it follows that

$$\phi(x,s) - \phi(x(\alpha), s(\alpha)) \geq -\frac{\alpha r_j}{| r_j |} v_j \|r\|_\infty - \frac{(\alpha v_j / v_{\min})^2}{2(1 - | \alpha | v_j / v_{\min})}.$$

Therefore, by choosing $\alpha = -\frac{|r_j|}{r_j} \frac{2v_{\min}/v_j}{3\sqrt{n+2}}$ and using Lemma 5.1, we obtain

$$\begin{aligned}
\phi(x,s) - \phi(x(\alpha), s(\alpha)) &> \frac{2}{3\sqrt{n(n+2)}} - \frac{2}{9(n+2)(1 - 2/(3\sqrt{n+2}))} \\
&> \frac{2}{3(n+2)}\left(1 - \frac{1}{2}\right) = \frac{1}{3(n+2)}.
\end{aligned}$$

□

Lemma 5.2 provides a possible step size rule for Step 3 in Algorithm 1. A better step size may be obtained by means of some line search procedure. Combining Lemma 4.1 and Lemma 5.2, we obtain the main result of this paper:

Theorem 5.1. *Algorithm 1 solves (P) and (D) in $\mathcal{O}(mn^{2.5}L)$ operations.*

Proof:

From Lemma 5.2 we know that in every iteration, the potential function is reduced by at least a quantity of $\frac{1}{3(n+2)}$. Using Lemma 4.1, this implies that Algorithm 1 can be used to solve (P) and (D) in $\mathcal{O}(n^{1.5}L)$ iterations. A single iteration involves $\mathcal{O}(mn)$ operations. Thus, in total the algorithm requires $\mathcal{O}(mn^{2.5}L)$ operations in the worst case.

$\square$

6. Concluding Remarks

It is known that interior point algorithms usually require significantly less iterations for solving linear programs than simplex algorithms. Yet, the simplex method can outperform the interior point method in running time on a wide range of problems because of the low computational complexity per step.

In this paper, we proposed an interior point algorithm where the amount of work per step is comparable to the work in a simplex iteration. As far as we know, this interesting property is not shared by any other interior point algorithm.

References

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

[2] Todd, M.J. and Ye, Y., "A centered projective algorithm for linear programming," *Mathematics of Operations Research* 15 (1990) 508-529.

[3] Ye, Y., "An $\mathcal{O}(n^3L)$ potential reduction algorithm for linear programming," *Mathematical Programming* 50 (1991) 239-258.