Yugoslav Journal of Operations Research
11 (2001), Number 1, 31-39

ON FINITE TERMINATION IN THE PRIMAL-DUAL METHOD FOR LINEAR PROGRAMMING

Neboj{a V. STOJKOVI]

Faculty of Economics, University of Ni{
Ni{, Yugislavia

nebojsas@orion.eknfak.ni.ac.yu

Abstract: In this paper we propose a modification of the finite termination algorithm which reduces the dimension of the primal-dual linear programming problem. We note that the similar approach is possible in any primal-dual algorithm for linear programming.

Keywords: Linear programming, interior-point methods, finite termination algorithm.

1. INTRODUCTION

We are concerned with the linear programming problem, which we write in the standard form as

$$\min c^T x \text{ subject to } Ax = b, \quad x \geq 0 , \tag{1.1}$$

where $c, x \in R^n$, $b \in R^m$, and A is an $m \times n$ real matrix. The dual problem for (1.1) is

$$\max b^T y \text{ subject to } A^T y + s = c, \quad s \geq 0 , \tag{1.2}$$

where $y \in R^m$ and $s \in R^n$.

It is known that the vector $x^* \in R^n$ is a solution of (1.1) if and only if there exist vectors $s^* \in R^n$ and $y \in R^m$ for which the following conditions hold:

$$A^T y^* + s* = c, \tag{1.3a}$$

$$Ax^* = b, \tag{1.3b}$$

$$x_i^* s_i^* = 0, \qquad i = 1, \dots, n, \tag{1.3c}$$

$$(x^*, s^*) \geq 0. \tag{1.3d}$$

All primal-dual methods generate iterates (x^k, y^k, s^k) that satisfy the bounds (1.3d) strictly, that is, $x^k > 0$ and $s^k > 0$, and instead condition (1.3c) deal with condition $x_i s_i = \tau, i = 1, \dots, n$, where $\tau \to 0$. We define the residuals for two linear equations as

$$r_b = Ax - b, \quad r_c = A^T y + s - c,$$

and use the following notations

$$X = \mathrm{diag}(x_1, \dots, x_n), \qquad S = \mathrm{diag}(s_1, \dots, s_n), \qquad e = (1, 1, \dots, 1)^T.$$

Let

$$\Omega_P = \{x^* \mid x^* \text{ solves } (1.1)\}, \qquad \Omega_D = \{(y^*, s^*) \mid (y^*, s^*) \text{ solves } (1.2)\}.$$

For every solution (x^*, y^*, s^*), we know that $x_j^* = 0$ and/or $s_j^* = 0$ for all $j = 1, \dots, n$. Let

$$B = \{ j \in \{1, \dots, n\} \mid x_j^* \neq 0 \text{ for some } x^* \in \Omega_P \},$$

$$N = \{ j \in \{1, \dots, n \mid s_j^* \neq 0 \text{ for some } (y^*, s^*) \in \Omega_D \}.$$

Clearly, $B \cap N = \varnothing$. Primal-dual strictly feasible set F^0 is

$$F^0 = \{(x, y, s) \mid Ax = b, \ A^T y + s = c, (x, s) > 0\}.$$

The next result is well known as the Goldman-Tucker theorem [5].

Definition 1.1. There exist at least one primal solution $x^* \in \Omega_P$ and one dual solution $(y^*, s^*) \in \Omega_D$ such that $x^* + s^* > 0$.

The solution from Theorem 1.1 is a strictly complementary solution. Note that if (x, y, x) and (x^*, y^*, s^*) are two strictly complementary solution pairs, then $x_i > 0$ if and only if $x_i^* > 0, i = 1, \dots, n$, and similarly, $s_i > 0$ if and only if $s_i^* > 0, i = 1, \dots, n$, i.e. the sets of indices of positive coordinates are the same for all strictly complementary optimal pairs [6].

Notice that in many practical applications of linear programming, a sequence of closely related problems has to be solved. When two closely related problems are solved problems the previous optimal solution should and can be used to solve the new problem faster. In the context of the simplex algorithm the aim is achieved by starting from the previous optimal basic solution. In the context of an interior-point method the warm start procedure still does exist [1], [4]. The approach adopted nowadays is to solve the first problem of a sequence of closely related problems using IPM and then cross-over to the simplex method. In this case the advantages of both methods are exploited. The algorithm to generate an optimal basis has been proposed by Megiddo [7]. It constructs an optimal basis in less than n iterations starting from any complementary solution, so Megiddo′s algorithm assumes that the exact optimal solution is known. This assumption is never encountered in practice, because the primal-dual algorithm only generates a sequence converging towards the optimal solution. Due to the finite precision of computations, the solution is neither exactly feasible nor complementary. The finite termination strategy proposed by Andersen and Ye [2], [10] attempts to jump from a path-following iterate to an exact primal-dual solution. The algorithm is sufficiently advanced - advanced enough for the index sets B and N to be well resolved. If not, the algorithm will produce a point that violates one or more constraints in (1.1), (1.2) and in this case we can return to the primal-dual method and take a few more steps before attempting finite termination again. In this paper we suggest an improvement of this algorithm.

1. FINITE TERMINATION ALGORITHM

In the sequel we consider the linear programming problem in the symmetric form. Note that the linear problem is usually given in that form and that every problem can be transformed in to the symmetric form. Consider the linear programming problem

$$\begin{aligned}
&\min\ c_1x_1+\cdots+c_lx_l \\
&\text{subject to} \\
&\quad a_{11}x_1+\cdots+a_{1l}x_l \le b_1 \\
&\qquad\qquad \cdots \\
&\quad a_{m1}x_1+\cdots+a_{ml}x_l \le b_m \\
&\quad x_i \ge 0, \quad i=1,\dots,l.
\end{aligned} \tag{2.1}$$

The standard form (2.1) is (with $l+m=n$)

$$
\begin{aligned}
&\min\ c_1x_1+\cdots+c_lx_l+\cdots+c_nx_n\\
&\text{subject to}\\
&\quad a_{11}x_1+\cdots+a_{1l}x_l+x_{l+1}=b_1\\
&\qquad\qquad \cdots\\
&\quad a_{m1}x_1+\cdots+a_{ml}x_l+x_{l+m}=b_m\\
&\quad x_i\ge 0,\quad i=1,\dots,n.
\end{aligned}
\tag{2.2}
$$

Notice that

$$c_{l+1}=\cdots c_{l+m}=0\ . \tag{2.3}$$

The dual problem (2.2) is

$$
\begin{aligned}
&\max\ b_1y_1+\cdots+b_my_m\\
&\text{subject to}\\
&\quad a_{11}y_1+\cdots+a_{m1}y_m+s_1=c_1\\
&\qquad\qquad \cdots\\
&\quad a_{1l}y_1+\cdots+a_{ml}y_l\ \ +s_l=c_l\\
&\quad\ \ y_1 \qquad\qquad\quad +s_{l+1}=c_{l+1}\\
&\qquad\qquad\quad y_m+s_{l+m}=c_{l+m}\\
&\quad s_i\ge 0,\quad i=1,\dots,n.
\end{aligned}
\tag{2.4}
$$

Let $x_{i_1}=x_{i_2}=\cdots=x_{i_p}=0$ and $I=\{i_1,\dots,i_p\}$. Denote by $\hat{A}$ the matrix A without i-th columns, $i\in I$ and let $\hat{x},\hat{c},\hat{s}\in R^{n-p}$ be, respectively, vectors x, c and s without i-th ($i\in I$) coordinates. Denote with $\varphi_i(y)$ linear function

$$\varphi_i(y)=c_i-(a_{1i}y_1+\cdots+a_{mi}y_m),\quad i\in I\ .$$

Lemma 2.1. Suppose that $x_{i_1}=x_{i_2}=\cdots=x_{i_p}=0$ is known. Then the primal-dual problem (2.2), (2.4) is equivalent to

$$\min\ \hat{c}^T\hat{x}\ \text{ subject to }\ \hat{A}\hat{x}=b,\quad \hat{x}\ge 0,\ x_i=0,\ i\in I\ ,$$

$$\max\ b^Ty\ \text{ subject to }\hat{A}y+\hat{s}=\hat{c},\quad \hat{s}\ge 0,\ s_i\in\varphi_i(y),\ i\in I\ .$$

Proof. It is enough to prove the Lemma for one index $i\in I$. Suppose now that $x_i=0$. We use the following notation:

$$\hat{A}^i=[A_1\cdots A_{i-1}A_{i+1}\cdots A_n]\ ,$$

$$\hat{x}^i=(x_1,\cdots,x_{i-1},x_{i+1},\cdots,x_n)\ ,$$

$$\hat{c}^i = (c_1, \cdots, c_{i-1}, c_{i+1}, \cdots, c_n),$$

$$\hat{s}^i = (s_1, \cdots, s_{i-1}, s_{i+1}, \cdots, s_n).$$

Now (2.2) is equivalent to $x_i = 0$ and

$$\min\ (\hat{c}^i)^T \hat{x}^i \text{ subject to } \hat{A}^i \hat{x}^i = b, \quad \hat{x}^i \geq 0. \tag{2.5}$$

The dual problem for (2.5) is

$$\max\ b^T y \text{ subject to } (\hat{A}^i)^T y + \hat{s}^i = \hat{c}^i, \quad \hat{s}^i \geq 0. \tag{2.6}$$

Now we have that (2.6) and

$$s_i = c_i - (a_{1i} y_1 + \cdots + a_{mi} y_m) = \varphi_i(y)$$

are equivalent to (2.4). ♦

Let $s_{l+j_1} = s_{l+j_2} = \cdots = s_{l+j_q} = 0$ and $J = \{ j_1, \ldots, j_q \}$. Denote with $\hat{A}^T$ the matrix A^T without j-th columns and without (l+j)-th rows, $j \in J$. Let $\hat{x}, \hat{c}, \hat{s} \in R^{n-q}$ be, respectively, vectors x, c, s without (l+j)-th ($j \in J$) coordinates, and let $\hat{y}, \hat{b} \in R^{m-q}$ be, respectively, vectors y, b without j-th ($j \in J$) coordinates. Denote with $\psi_j(\hat{x})$ linear function

$$y_i(\hat{x}) = b_j - (a_{j1} x_1 + \cdots + a_{jl} x_l).$$

Lemma 2.2. Suppose that $s_{l+j_1} = s_{l+j_2} = \cdots = s_{l+j_q} = 0$ is known. Then the primal-dual problem (2.2), (2.4) is equivalent to

$$\min \hat{c}^T \hat{x} \text{ subject to } \hat{A}\hat{x} = \hat{b}, \quad \hat{x} \geq 0, \ x_{l+j} = \psi_j(\hat{x}), \ j \in J,$$

$$\max \hat{b}^T \hat{y} \text{ subject to } \hat{A}^T \hat{y} + \hat{s} = \hat{c}, \quad \hat{s} \geq 0, \ s_j = 0, \ j \in J.$$

Proof. It is enough to prove the Lemma for one index $j \in J$. Suppose now that $s_{l+j} = 0$ for some $1 \leq j \leq m$. Because (2.3) we have $y_j = 0$, and (2.4) is equivalent to

$$\max \hat{b}^T \hat{y} \text{ subject to } \hat{A}^T \hat{y} + \hat{s} = \hat{c}, \quad \hat{s} \geq 0, \tag{2.7}$$

where $(\hat{A}^j)^T$ is a matrix A^T without j-th columns and without (l+j)-th rows and

$$\hat{y}^j = (y_1, \cdots, y_{j-1}, y_{j+1}, \cdots, y_m),$$

$$\hat{b}^j = (b_1, \cdots, b_{j-1}, b_{j+1}, \cdots, b_m),$$

$$\hat{s}^j = (s_1, \cdots, s_{l+j-1}, s_{l+j+1}, \cdots, s_n),$$

$$\hat{c}^j = (c_1, \cdots, c_{l+j-1}, c_{l+j+1}, \cdots, c_n).$$

The primal problem for (2.7) is

$$\min(\hat{c}^j)^T \hat{x}^j \text{ subject to } \hat{A}^j \hat{x}^j = \hat{b}, \quad \hat{x}^j \geq 0, \tag{2.8}$$

where

$$\hat{x}^j = (x_1, \cdots, x_{l+j-1}, x_{l+j+1}, \cdots, x_n).$$

As (2.2) is equivalent to (2.8) with

$$x_{j+l} = b_j - (a_{j1}x_1 + \cdots + a_{jl}x_l) = \psi_j(\hat{x}),$$

the proof follows. ♦

Note that, because Theorem 1.1, $I \cap J^l = \varnothing$, $J^l = \{ j+l \mid j \in J\}$ for strictly complementary solution. The next theorem immediately follows from Lemma 2.1 and Lemma 2.2.

Theorem 2.3. Suppose that $x_{i_1} = x_{i_2} = \cdots = x_{i_p} = 0$ and $s_{l+j_1} = s_{l+j_2} = \cdots = s_{l+j_q} = 0$ is known. Then the primal-dual problem (2.2), (2.4) is equivalent to

$$\min\ \hat{c}^T \hat{x} \text{ subject to } \hat{A}\hat{x} = \hat{b},$$

$$\hat{x} \geq 0,\ x_i = 0, i \in I, x_{l+j} = \psi_j(\hat{x}),\ j \in J,$$

$$\max\ \hat{b}^T \hat{y} \text{ subject to } \hat{A}^T \hat{y} + \hat{s} = \hat{c},$$

$$\hat{s} \geq 0,\ s_j = 0,\ j \in J, s_i = j_i(\hat{y}),\ i \in I.$$

A theorem similar to Theorem 2.3 is proved in [8]. It is evident that Theorem 2.3 is valid even in the case when in (2.1) we have some equalities.

Note that for any point $(x, y, s) \in N_{-\infty}(\gamma)$, we can estimate the set B and N as follows [8]:

$$\begin{aligned} B(x,s) &= \{i \in \{1,\ldots,n\} \mid x_i \geq s_i\}, \\ N(x,s) &= \{1,\ldots,n\} \setminus B(x,s), \end{aligned} \tag{2.9}$$

where

$$N_{-\infty}(\gamma) = \{(x, y, s) \in F^0 \mid x_i s_i \geq \gamma\mu \quad \text{for all } i = 1,\ldots,n\}$$

where $\gamma = (0,1)$.

The next finite termination is due to Ye [8].

Procedure FT

Given $\gamma = (0,1)$ and $(x, y, s) \in N_{-\infty}(\gamma)$:

Find $B(x,s)$ and $N(x,s)$ from (2.9);

Solve the following problem:

$$\min_{(x^*,y^*,s^*)} \frac{1}{2}\left\|x^* - x\right\|^2 + \frac{1}{2}\left\|s^* - s\right\|^2, \tag{2.10a}$$

$$Ax^* = b, \quad A^T y^* + s^* = c, \tag{2.10b}$$

$$x^* = 0 \text{ for } i \in N(x,s), \quad s_i^* = 0 \text{ for } i \in B(x,s), \tag{2.10c}$$

if $x_B^* > 0$ and $s_N^* > 0$

declare success: (x^*, y^*, s^*) is a strictly complementary solution;

else

declare failure and return to the primal-dual algorithm.

The following result [9] shows that a successful outcome for Procedure FT is guaranteed when μ is sufficiently small.

Theorem 2.4. Let $\gamma = (0,1)$ be given. Then there is a threshold value $\bar{\mu}$ such that for all (x, y, s) that satisfy

$$(x, y, s) \in N_{-\infty}(\gamma), \quad 0 < \mu = x^T s / n \le \bar{\mu},$$

we have

i) $B(x,s) = B$ and $N(x,s) = N$; that is, actual index sets B and N are estimated correctly by the procedure (2.9);

ii) the projection procedure (2.10) yields a strictly complementary solution (x^*, y^*, s^*).

An important practical issue is the choice of indicator $B(x,s)$ for optimal partition B. Indicator (2.9) is not invariant with respect to the column scaling. A better indicator is [3].

$$B(x,s)=\{i\in\{1,\dots,n\}\,|\,(|\Delta x_i^{\alpha}|/x_i)\le(|\Delta s_i^{\alpha}|/s_i)\}\,, \tag{2.11}$$

where $(\Delta x_i^{\alpha},\Delta s_i^{\alpha})$ is primal-dual affine scaling search direction. This indicator is scaling invariant [1].

Now we propose the following modification of the finite termination procedure. Let $\varepsilon>0$ be given. Use (2.9) or (2.11) to estimate sets B and N. Sets

$$I=\{i\,|\,x_i<\varepsilon, i\in N\}=\{i_1,\dots,i_p\}\,,$$

$$J=\{j\,|\,s_{l+j}<\varepsilon, l+j\in B\}=\{j_1,\dots,j_p\}\,.$$

in Theorem 2.3 and if

$$x_{l+j}=y_j(\hat{x})>0,\ j\in J,\text{and } s_i=j_i(\hat{y})>0, i\in I\,, \tag{2.12}$$

apply Procedure FT on the primal-dual problem

$$\min\ \hat{c}^T\hat{x}\ \text{subject to}\ \hat{A}\hat{x}=\hat{b}\,,$$

$$\max\ \hat{b}^T\hat{y}\ \text{subject to}\ \hat{A}^T\hat{y}=\hat{c}\,.$$

where $\hat{A}$ is matrix without i-th and $(l+j)$-th columns, and without j-th rows, $i\in I=\{i_1,\dots,i_p\}$, $j\in J=\{j_1,\dots,j_q\}$. Condition (2.12) ensures that we have a strictly complementary solution. Note that in this variant of algorithm we reduce the primal-dual problem with matrix $A_{m\times n}$ to the primal-dual problem with matrix $\hat{A}_{(m-q)\times(n-q-p)}$.

A similar approach is possible in the iteration steps of any primal-dual algorithm. It is known that the linear system to be solved at each primal-dual iteration can be formulated in three equivalent ways. The unreduced form for the infeasible-interior-point algorithm is

$$\begin{bmatrix} 0 & A & 0\\ A^T & 0 & I\\ 0 & S & X\end{bmatrix}\begin{bmatrix}\Delta y\\ \Delta x\\ \Delta s\end{bmatrix}=\begin{bmatrix}-r_b\\ -r_c\\ -XS_e+\sigma\mu e\end{bmatrix}, \tag{2.13}$$

where $\mu=x^T s/n$. Eliminating Δs form (2.13) and using notation $D=S^{-1/2}X^{1/2}$, we obtain the augmented system

$$\begin{bmatrix} 0 & A\\ A^T & -D^{-2}\end{bmatrix}\begin{bmatrix}\Delta y\\ \Delta x\end{bmatrix}=\begin{bmatrix}-r_b\\ -r_c+s-\sigma\mu X^{-1}e\end{bmatrix}, \tag{2.14a}$$

$$\Delta s=-s+\sigma\mu X^{-1}e-X^{-1}S\Delta x\,. \tag{2.14b}$$

Now, we can eliminate Δx from (2.14a) to obtain the normal equations form

$$AD^2A^T\Delta y = -r_b - A(S^{-1}Xr_c + x - \sigma\mu S^{-1}e)\ , \tag{2.15a}$$

$$\Delta s = -r_c - A^T\Delta y\ , \tag{2.15b}$$

$$\Delta x = -x + \sigma\mu S^{-1}e + S^{-1}X\Delta s)\ . \tag{2.15c}$$

The normal equations form is used by most primal-dual codes. For both (2.14) and (2.15), particular issues of stability arise because the presence of very small and very large diagonal elements is both

$$D^2 = \mathrm{diag}(x_1/s_1,\dots,x_n/s_n) \text{ and } D^{-2} = \mathrm{diag}(s_1/x_1,\dots,s_n/x_n)\ .$$

Applying Theorem 2.3 we can eliminate some of these elements and improve stability and also improve the centrality of the iteration sequence.

Acknowledgemts

I am grateful to Prof. P. Staminirovi} for his very detailed comments on the paper.

REFERENCES

[1] Andersen, E.D., Gondzio, J., MõszÃros, C., and Xu, X., "Implementation of interior point methods for large scale linear programming", Technical Report, HEC, UniversitÕ de GenÕve, 1996.

[2] Andersen, E.D., and Ye, Y., "Combining interior-point and pivoting algorithms for linear programming", Technical Report, Department of Management Sciences, The University of Iowa, 1994

[3] El-Barky, A.S., Tapis, R.A. and Zhang ,Y., "A study of indicators for identifying zero variables in interior-point methods", SIAM Rcv, 36 (1) (1994) 45-72.

[4] Freund, R. M., "A potential-function reduction algorithm for solving a linear program directly from an infeasible 'warm start'", Mathematical Programming, 52 (1991) 441-446.

[5] Goldman, A.J., and Tucker A.W., "Theory of linear programming", in: H W. Kuhn and A.W. Tucker (eds.), Linear Equalities and Related Systems, Princeton University Press, Princeton, N.J., 1956, 53-97.

[6] Kova~evi}-Vuj~i}, V. and A{i}, M.D., "Stabilization of interior-point methods for linear programming", Computational Optimization and Applications, 14 (1999) 1-16.

[7] Megiddo N., "On finding primal- and dual-optimal bases", ORSA Journal on Computing, 3 (1991) 63-65.

[8] Stojkovi}, N.V., and Stanimirovi}, P.S., "On the elimination of excessive constraints in linear programming", in: M. Vujo{evi} and M. Marti} (eds.), Proceedings of XXVI Yugoslav Symposium on Operations Research, Belgrade, 1999, 207-210.

[9] Wright, S. J., Primal-dual Interior-Point Methods, SIAM, Philadelphia, 1997.

[10] Ye Y., "On the finite convergence of interior-point algorithms for linear programming", Mathematical Programming, 57 (1992) 325-336.