TRANSFORMATIONS OF DUAL PROBLEM AND DECREASING DIMENSIONS IN LINEAR PROGRAMMING

Nebojša V. Stojković and Predrag S. Stanimirović

Abstract. We investigate behavior of the potential function in a modification of the Mehrotra's primal-dual algorithm. This modification reduces dimensions of the problem and eliminates need for the finite termination algorithm. Numerical results on some examples from the Netlib test set are provided.

We also regard problems about applying a stabilization procedure proposed by Kovačević-Vujčić and Ašić in the Mehrotra's primal dual interior-point algorithm for linear programming. Transformations of the dual problem required for the application of the stabilization procedure are considered.

1. Introduction

Primal-dual algorithms investigate linear programming problem in the standard form as

minimize
$$c^T x$$
 subject to $Ax = b$, $x > 0$, (1.1)

where $c, x \in \mathbb{R}^n$, $b \in \mathbb{R}^m$, A is an $m \times n$ real matrix and c^T is transpose of the vector c. The dual problem for (1.1) is

maximize
$$b^T \lambda$$
 subject to $A^T \lambda + s = c, \quad s \ge 0,$ (1.2)

where $\lambda \in \mathbb{R}^m$ and $s \in \mathbb{R}^n$ and b^T , A^T denote transpose of the vector b and the matrix A, respectively. It is known that the vector $x^* \in \mathbb{R}^n$ is a solution of (1.1) if and only if there exist vectors $s^* \in \mathbb{R}^n$ and $\lambda^* \in \mathbb{R}^m$ such that the following conditions hold:

$$A^T \lambda^* + s^* = c, \qquad Ax^* = b,$$
 (1.3a-b)

$$x_i^* s_i^* = 0, \quad i = 1, \dots, n, \qquad (x^*, s^*) \ge 0.$$
 (1.3c-d)

The central path $\mathcal C$ is an arc of strictly feasible points that plays a vital role in the theory of primal-dual algorithms. It is parametrized by a scalar $\tau > 0$, such

AMS Subject Classification: 90 C 05

Keywords and phrases: Primal-dual interior point methods, potential function.

Communicated at the 5th International Symposium on Mathematical Analysis and its Applications, Niška banja, Yugoslavia, October, 2–6, 2002.

that each point $(x_{\tau}, \lambda_{\tau}, s_{\tau}) \in \mathcal{C}$ solves the following system:

$$A^T \lambda + s = c, \qquad Ax = b, \tag{1.4a-b}$$

$$x_i s_i = \tau, \quad i = 1, \dots, n, \qquad (x, s) \ge 0.$$
 (1.4c-d)

Instead of the complementarity conditions (1.3c), in (1.4) it is required that the pairwise products $x_i s_i$ have the same value for all indices i. From (1.4) we can define the central path as

$$\mathcal{C} = \{(x_{\tau}, \lambda_{\tau}, s_{\tau}) \mid \tau > 0\}.$$

The primal-dual feasible set \mathcal{F} and strictly feasible set \mathcal{F}^0 are defined by (see [7])

$$\mathcal{F} = \{ (x, \lambda, s) \mid Ax = b, \ A^T \lambda + s = c, \ (x, s) \ge 0 \},$$

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

The paper is organized as follows. In the second section we consider behavior of the potential-reduction function caused by the elimination of columns and rows of the matrix A, corresponding to zero variables. This part of the paper is a continuation of the paper [6]. The motivation for this section is based on the fact that reducing dimensions of the problem improves the stability and centrality of iterative sequence. We show that reducing dimension decrease the potential function Φ_{ρ} . In the third section we investigate transformations of the dual problem caused by the stabilization procedure introduced in [3]. In the last section we show the decreasing of the potential-reduction function in several known test problems.

2. Reduction of linear programming problem and potential function

In sequel we regard the linear programming problem in the general form

minimize
$$c_1x_1 + \cdots + c_kx_k$$

subject to
$$\sum_{j=1}^{k} a_{ij} x_{j} \leq b_{i}, \quad i = 1, ..., q,$$

$$\sum_{j=1}^{k} a_{ij} x_{j} = b_{i}, \quad i = q + 1, ..., m,$$
(2.1)

$$x_j \ge 0, \quad j = 1, \dots, k.$$

Standard form of (2.1) is (with k + q = n)

minimize
$$c_1 x_1 + \cdots + c_k x_k + c_{k+1} x_{k+1} + \cdots + c_n x_n$$

subject to
$$\sum_{j=1}^{k} a_{ij}x_j + x_{k+i} = b_i, \quad i = 1, \dots, q,$$
$$\sum_{j=1}^{k} a_{ij}x_j = b_i, \quad i = q+1, \dots, m,$$
 (2.2)

$$x_j \geq 0, \quad j = 1, \dots, n.$$

The dual problem for (2.2) is

maximize
$$b_1\lambda_1 + \cdots + b_m\lambda_m$$

subject to
$$\sum_{j=1}^{m} a_{ji}\lambda_{j} + s_{i} = c_{i}, i = 1, ..., k$$

 $\lambda_{j} + s_{k+j} = 0, j = 1, ..., q$
 $s_{i} > 0, i = 1, ..., k + q.$ (2.3)

It is clear that if $x_i=0$ for some i, then this x_i has no influence on the final solution, so i-th column from the matrix A can be omitted. Also, if $s_{k+j}=0$ for some j, then from (2.3) it follows $\lambda_j=0$, so jth column and (k+j)-th row from the matrix A^T can be omitted. These facts imply the next two statements from [6]. For the sake of simplicity we use the following notations. Let $\alpha=\{\alpha_1,\ldots,\alpha_s\}$ and $\beta=\{\beta_1,\ldots,\beta_t\}$ be subsets of $\{1,\ldots,m\}$ and $\{1,\ldots,n\}$, respectively, for some integers $1\leq s\leq m$ and $1\leq t\leq n$. By A^α we denote the $p\times n$ submatrix of A determined by the entries in rows indexed by α . Similarly, by A_β we denote the $m\times q$ submatrix of A determined by the entries in columns indexed by β .

Lemma 2.1. [6] Consider the linear programming problem (2.2),(2.3). Assume that $x_{i_1} = \cdots = x_{i_d} = 0$, and consider the set of indices $I = \{i_1, \ldots, i_d\} \subseteq \{1, \ldots, n\}$. Denote by $\hat{A} \in \mathbb{R}^{m \times (n-d)}$ the matrix $A_{\{1,\ldots,n\}\setminus I}$, and let $\hat{x}, \hat{c}, \hat{s} \in \mathbb{R}^{n-d}$ be, respectively, vectors x, c and s without i-th $(i \in I)$ coordinates $(\hat{x} = x_{\{1,\ldots,n\}\setminus I}, \hat{c} = c_{\{1,\ldots,n\}\setminus I}$ and $\hat{s} = s_{\{1,\ldots,n\}\setminus I}$. By $\varphi_i(\lambda)$ we denote the linear function

$$\varphi_i(\lambda) = c_i - (a_{1i}\lambda_1 + \dots + a_{mi}\lambda_m), \quad i \in I.$$

Then the primal-dual problem (2.2), (2.3) is equivalent to

minimize $\hat{c}^T \hat{x}$ subject to $\hat{A}\hat{x} = b$, $\hat{x} \geq 0$, $x_i = 0$, $i \in I$,

maximize $b^T \lambda$ subject to $\hat{A}^T \lambda + \hat{s} = \hat{c}$, $\hat{s} > 0$, $s_i = \varphi_i(\lambda)$, $i \in I$,

where \hat{A}^T is the transpose of \hat{A} .

LEMMA 2.2. [6] Assume that in the primal-dual problem (2.2),(2.3) the identities $s_{k+j_1} = \cdots = s_{k+j_h} = 0$ are satisfied. Consider the sets $J = \{j_1, \ldots, j_h\}$ and $J^k = \{k+j | j \in J\}$. Denote by $\hat{A}^T \in \mathbb{R}^{(n-h)\times(m-h)}$ the matrix

$$(A^T)_{\{1,\ldots,m\}\backslash J}^{\{1,\ldots,n\}\backslash J^k} = \left(A_{\{1,\ldots,n\}\backslash J^k}^{\{1,\ldots,m\}\backslash J}\right)^T.$$

Let $\hat{x}, \hat{c}, \hat{s} \in \mathbb{R}^{n-h}$ be, respectively, vectors x, c, s without (k+j)th $(j \in J)$ coordinates, and let $\hat{\lambda}, \hat{b} \in \mathbb{R}^{m-h}$ be, respectively, vectors λ, b without jth $(j \in J)$ coordinates. By $\psi_j(\hat{x})$ denote the linear function

$$\psi_j(\hat{x}) = b_j - \sum_{i=1, i \notin J^k}^n a_{ji} x_i, \quad j \in J.$$

Then the primal-dual problem (2.2),(2.3) is equivalent to

minimize $\hat{c}^T \hat{x}$ subject to $\hat{A}\hat{x} = \hat{b}$, $\hat{x} \ge 0$, $x_{k+j} = \psi_j(\hat{x})$, $j \in J$,

maximize $\hat{b}^T \hat{\lambda}$ subject to $\hat{A}^T \hat{\lambda} + \hat{s} = \hat{c}$, $\hat{s} \geq 0$, $s_{k+j} = 0$, $j \in J$,

where $\hat{A} = A_{\{1,\ldots,n\}\backslash J^k}^{\{1,\ldots,m\}\backslash J}$.

The following statement summarizes the results of above lemmas.

THEOREM 2.1. [6] Let $x_{i_1} = \cdots = x_{i_d} = 0$ and $s_{k+j_1} = \cdots = s_{k+j_h} = 0$, where the sets $I = \{i_1, \ldots, i_d\}$ and $J = \{j_1, \ldots, j_h\}$ are defined as in Lemma 2.1 and Lemma 2.2, respectively. Under the notations of Lemma 2.1 and Lemma 2.2 the primal-dual problem (2.2)-(2.3) is equivalent to

minimize $\hat{c}^T\hat{x}$ subject to $\hat{A}\hat{x} = \hat{b}$, $\hat{x} \ge 0$, $x_i = 0$, $i \in I$, $x_{k+j} = \psi_j(\hat{x})$, $j \in J$,

 $\begin{aligned} & \textit{maximize} \quad \hat{b}^T \hat{\lambda} \quad \textit{subject to} \quad \hat{A}^T \hat{\lambda} + \hat{s} = \hat{c}, \ \hat{s} \geq 0, \ s_{k+j} = 0, \ j \in J, \ s_i = \varphi_i(\hat{\lambda}), \ i \in I, \\ & \textit{where } \hat{A} \in \mathbb{R}^{(m-h) \times (n-h-d)}, \ \hat{x}, \hat{c}, \hat{s} \in \mathbb{R}^{n-h-d}, \ \hat{\lambda}, \hat{b} \in \mathbb{R}^{m-h}. \end{aligned}$

Remark 2.1. Theorem 2.1 is a generalization of Theorem 3.1, part (b), proved in [5].

The ability to identify zero variables early on in an iterative method is of considerable value and can be used to computational advantage. The paper [2] gives a formal presentation of the notion of indicators for identifying zero variables, and also studies various indicators. About the numerical experience with various indicators for identifying zero variables see also the paper [4]. The term indicator denotes a function that identifies constraints that are active at a solution of a constrained optimization problem. If some members of a solution x^* are identified early on in an iterative procedure, this information can be used as follows:

- 1. Reduce dimension of the problem by dropping the columns od A corresponding to zero variables.
 - 2. Help recover an optimal basis for the linear program.
 - 3. Help obtain very accurate solutions.

The reduction of dimension of the problem by dropping the columns of the matrix A corresponding to zero variables is investigated in [2] and [4]. We consider the influence of the elimination of columns and rows in the decreasing the potential function Φ_{ρ} .

Potential-reduction methods use a logarithmic potential function to measure the worth of each point in \mathcal{F}^0 and aim to achieve a certain fixed reduction in this function at each iteration. The most interesting primal-dual potential function is Tanabe-Todd-Ye function Φ_{ρ} , defined by

$$\Phi_{\rho}(x,s) = \rho \log x^{T} s - \sum_{i=1}^{n} \log x_{i} s_{i}$$

for some parameter $\rho > n$. In the next theorem we investigate the influence of reductions enabled by Theorem 2.1 to the function Φ_{ρ} .

Theorem 2.2. Suppose that the central path is not followed and suppose that $x_js_j<\varepsilon<1$ for some $j\in J\subset\{1,\ldots,n\}$ and $x_is_i\geq\varepsilon,$ $i\notin J$. If it is possible to eliminate x_j (or s_j , for some j>k) and to eliminate j-th columns from the matrix A (j-th rows and (j-k)-th columns from the matrix A^T), then

$$\log\left(1 + \frac{\sum_{j \in J} x_j s_j}{\sum_{i \notin J} x_i s_i}\right)^{\rho} - |J| \log \varepsilon \le \Phi_{\rho}(x, s) - \hat{\Phi}_{\rho}(\hat{x}, \hat{s})$$

$$\le \log\left(1 + \frac{|J|}{n - |J|}\right)^{\rho} - \sum_{j \in J} \log x_j s_j.$$

Proof. Applying Lemma 2.1 (Lemma 2.2) we can eliminate x_j , (s_j) and j-th columns from the matrix A (j-th rows and (j-k)th columns from the matrix A^T). Let $\hat{\Phi}_{\rho}(\hat{x},\hat{s}) = \rho \log \hat{x}^T \hat{s} - \sum_{i \notin J} \log x_i s_i$ be the potential-reduction function of reduced problem, where \hat{x} and \hat{s} denote the vectors x and s without j-th elements, respectively. Then we have

$$\begin{split} \Phi_{\rho}(x,s) - \hat{\Phi}_{\rho}(\hat{x},\hat{s}) &= \rho(\log x^T s - \log \hat{x}^T \hat{s}) - \sum_{j \in J} \log x_j s_j \\ &= \rho \log \left(1 + \frac{\sum_{j \in J} x_j s_j}{\sum_{i \notin J} x_i s_i} \right) - \sum_{j \in J} \log x_j s_j \\ &\geq \rho \log \left(1 + \frac{\sum_{j \in J} x_j s_j}{\sum_{i \notin J} x_i s_i} \right) - |J| \log \varepsilon. \end{split}$$

On the other side, since

$$\rho \log \left(1 + \frac{\sum_{j \in J} x_j s_j}{\sum_{i \notin J} x_i s_i}\right) \leq \rho \log \left(1 + \frac{|J|\varepsilon}{(n - |J|)\varepsilon}\right)$$

we have

$$\Phi_{\rho}(x,s) - \hat{\Phi}_{\rho}(\hat{x},\hat{s}) \le \log\left(1 + \frac{|J|}{n - |J|}\right)^{\rho} - \sum_{i \in J} \log x_i s_i,$$

which completes the proof. ■

3. Transformation of dual problem

The next procedure is proposed in [3] for stabilization of ill-conditioned linear programming problem. This stabilization procedure is based on the Gaussian elimination.

PROCEDURE 3.1. Step 1. For a given k let $j(1),\ldots,j(n)$ be the indices satisfying conditions $(x_{j(1)}^k/s_{j(1)}^k) \geq \cdots \geq (x_{j(n)}^k/s_{j(n)}^k)$, and let $\hat{h} \in \{1,\ldots,n-1\}$ be the smallest index such that $(x_{j(\hat{h}+1)}^k/s_{j(\hat{h}+1)}^k) < 1$.

Step 2. Order the columns of A and c^T so that the j(p)-th column comes to the p-th position, p = 1, ..., n. Set p = 1, q = 1.

Step 3. Find $|a_{i(p),p}| = \max\{|a_{ip}|, i = q, \dots, m\}$. If $|a_{i(p),p}| \neq 0$ go to Step 5.

Step 4. If $|a_{i(p),p}|=0$ and $p=\hat{h}$ set $\hat{r}=q-1$ and RETURN. Otherwise replace p by p+1 and go to Step 3.

Step 5. Exchange the i(p)-th and q-th row in A and b. Using a_{qp} as the pivot element eliminate a_{ip} , $i = q + 1, \ldots, m$ (if any) and c_p .

Step 6. If $q = \min\{\hat{h}, m\}$ or $p = \hat{h}$ set $\hat{r} = q$ and RETURN. Otherwise replace p by p + 1, q by q + 1 and go to Step 3.

In order to incorporate this procedure in primal-dual interior point method, we need corresponding modification of Step~5 in the stabilization procedure. This modification gives a specific transformation of the dual problem. We regard linear problem (1.1) with objective functions $c^Tx + d$ and $b^T\lambda + d$ for primal and dual problem respectively.

Theorem 3.1. Suppose that the matrix A of linear problem (1.1), which satisfies $c_1 = \cdots = c_{p-1} = 0$, has a form

$$A = \begin{bmatrix} a_{11} & \dots & a_{1,p-1} & a_{1p} & \dots & a_{1n} \\ \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & a_{q-1,p-1} & a_{q-1,p} & \dots & a_{q-1,n} \\ 0 & \dots & 0 & a_{qp} & \dots & a_{qn} \\ \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & 0 & a_{mp} & \dots & a_{mn} \end{bmatrix} .$$
 (3.1)

If we use a_{qp} as the pivot element to eliminate $a_{ip}, i = q + 1, ..., m$, and c_p then the dual problem of transformed problem

minimize
$$c'^T x + d'$$
 subject to $A' x = b'$, $x \ge 0$,

is

maximize
$$b^{\prime T} \lambda^{\prime} + d^{\prime}$$
 subject to $A^{\prime T} \lambda^{\prime} + s = c^{\prime}$, $s > 0$,

where

$$d' = d + \frac{c_p}{a_{qp}}b_q, \quad \lambda_i' = \lambda_i, \quad i \neq q, \quad \lambda_q' = \lambda_q + \sum_{i=q+1}^m \frac{a_{ip}}{a_{qp}}\lambda_i - \frac{c_p}{a_{qp}}.$$

Proof. As

$$b_i' = b_i, \quad i = 1, \dots, q, \qquad b_i' = b_i - \frac{a_{ip}}{a_{qp}} b_q, \quad i = q + 1, \dots, m,$$

the value of dual objective function is

$$b'^{T}\lambda' + d' = \sum_{i=1}^{m} b'_{i}\lambda'_{i} + d'$$

$$= \sum_{i=1}^{q-1} b_{i}\lambda_{i} + b_{q} \left(\lambda_{q} + \sum_{i=q+1}^{m} \frac{a_{ip}}{a_{qp}}\lambda_{i} - \frac{c_{p}}{a_{qp}}\right) + \sum_{i=q+1}^{m} \left(b_{i} - \frac{a_{ip}}{a_{qp}}b_{q}\right)\lambda_{i} + d'$$

$$= \sum_{i=1}^{q} b_{i}\lambda_{i} + \sum_{i=q+1}^{m} \frac{a_{ip}}{a_{qp}}b_{q}\lambda_{i} - \frac{c_{p}}{a_{qp}}b_{q} + \sum_{i=q+1}^{m} b_{i}\lambda_{i} - \sum_{i=q+1}^{m} \frac{a_{ip}}{a_{qp}}b_{q}\lambda_{i} + d'$$

$$= b^{T}\lambda - \frac{c_{p}}{a_{qp}}b_{q} + d' = b^{T}\lambda + d$$

Further, the elements of matrix $A'^T \lambda'$ have the form $\sum_{i=1}^m a'_{ij} \lambda'_i$, $j = 1, \ldots, n$. Now we have

$$\begin{split} \sum_{i=1}^{m} a'_{ij} \lambda'_i + s_j &= \sum_{i=1}^{m} a_{ij} \lambda_i + s_j = 0 = c'_j, \quad j = 1, \dots, p-1, \\ \sum_{i=1}^{m} a'_{ip} \lambda'_i + s_p &= \sum_{i=1}^{q-1} a_{ip} \lambda_i + a_{qp} \Big(\lambda_q + \sum_{i=q+1}^{m} \frac{a_{ip}}{a_{qp}} \lambda_i - \frac{c_p}{a_{qp}} \Big) + s_p \\ &= \sum_{i=1}^{m} a_{ip} \lambda_i + s_p - c_p = 0 = c'_p, \\ \sum_{i=1}^{m} a'_{ij} \lambda'_i + s_j &= \sum_{i=1}^{q-1} a_{ij} \lambda_i + a_{qj} \Big(\lambda_q + \sum_{i=q+1}^{m} \frac{a_{ip}}{a_{qp}} \lambda_i - \frac{c_p}{a_{qp}} \Big) + \sum_{i=q+1}^{m} \Big(a_{ij} - \frac{a_{qj}}{a_{qp}} a_{ip} \Big) \lambda_i \\ &+ s_j &= \sum_{i=1}^{m} a_{ij} \lambda_i + s_j - \frac{a_{qj}}{a_{qp}} c_p = c_j - \frac{a_{qj}}{a_{qp}} c_j = c'_j, \quad j = p+1, \dots, n. \quad \blacksquare \end{split}$$

Remark 3.1. It is obvious that the form (3.1) of the matrix A is not restrictive because, if a_{11} is the pivot element for arbitrary matrix A and arbitrary c, we immediately get the matrix of the form (3.1) and after applying the transformation, matrix A' has the same form.

4. Numerical experiences

The reduction of dimensions in Mehrotra's algorithm is implemented in the package MATHEMATICA. In the next examples we consider a small subset of known test problems in the literature. We use notation $delta = \Phi_{\rho}(x,s) - \hat{\Phi}_{\rho}(\hat{x},\hat{s})$. Note that practical value of ρ are $n + \sqrt{n}$, 10n or $n + n^{1.5}$. In the following examples we used compromise value $\rho = 2n$.

Example 4.1. For the test problem Afiro we get the result arranged in the Table. Eliminating threshold is 10^{-1} . Note that in 8th iteration 35 variables are eliminated and we have considerable decreasing of the potential function. Column *No iter*. denotes the order number of the iteration when the elimination is applied, and in column *No var*. we give the number of eliminated variables.

Afiro

EXAMPLE 4.2. In the tables on the next page we give results for some *Netlib* test problems. We see that the potential-reduction function is considerably decreased when the elimination is applied.

5. Concluding remarks

We investigate the elimination of columns and constraints in the primal-dual interior point method which is mainly based on the Mehrotra's algorithm, restated in [7], [1]. It is well known that reducing dimensions of the problem improve the stability and centrality of iterative sequence. We showed that reducing dimension decrease the potential function. In order to make numerical experience, we develop a research code for the implementation of the Mehrotra's algorithm and the modification in the package MATHEMATICA.

We also investigate transformations of the dual problem, caused by *Step 5* of the stabilization procedure introduced in [3].

REFERENCES

- [1] E.D. Andersen, J. Gondzio, C. Mészáros and X. Xu, *Implementation of interior point methods for large scale linear programming*, Technical report, HEC, Université de Genève, 1996.
- [2] A.S. El-Bakry, R.A.Tapia, and Y. Zhang, A study of indicators for identifying zero variables in interior-point methods, SIAM Rev. 36(1) (1994), 45-72.
- [3] V.V. Kovačević-Vujčić and M.D. Ašić, Stabilization of interior-point methods for linear programming, Computational Optimization and Applications 14 (1999), 1-16.
- [4] I.J. Lustig, R.E. Marsten and D.F. Shanno, Computational experience with a primal-dual interior point method for linear programming, Linear Algebra Appl., 152 (1991), 191-222.
- [5] N.V. Stojković, and P.S. Stanimirović, Two direct method in linear programming, EJOR, 131 (2001), 417-439.

- [6] N.V. Stojković, On the finite termination in the primal-dual algorithm for linear programming, YUJOR, ${\bf 11}~(2001),~31-40.$
- $\label{eq:continuous} \textbf{[7]} \ \ \text{S.J. Wright}, \ Primal-dual \ Interior-point \ Methods}, \ \text{SIAM}, \ Philadelphia, \ 1997.$

(received 19.12.2002)

Faculty of Economics, Tr
g Kralja Aleksandra 11, 18000 Niš, Yugoslavia, $\emph{E-mail}\colon$ nebojsas@orion.ekn
fak.ni.ac.yu

University of Niš, Faculty of Science, Department of Mathematics, Ćirila i Metodija 2, 18000 Niš, Yugoslavia, E-mail: pecko@pmf.pmf.ni.ac.yu