

ON THE EXISTENCE OF A SHORT PIVOTING SEQUENCE FOR A LINEAR PROGRAM

Anders FORSGREN* Fei WANG†

August 29, 2019

Abstract

Pivoting methods are of vital importance for linear programming, the simplex method being the by far most well-known. In this paper, a primal-dual pair of linear programs in canonical form is considered. We show that there exists a sequence of pivots, whose length is bounded by the minimum dimension of the constraint matrix, such that the pivot creates a nonsingular submatrix of the constraint matrix which increases by one row and one column at each iteration. Solving a pair of linear equations for each of these submatrices generates a sequence of optimal solutions of a primal-dual pair of linear programs of increasing dimensions, originating at the origin. The optimal solutions to the original primal-dual pair of linear programs are obtained in the final step.

It is only an existence result, we have not been able to generate any rules based on properties of the problem to generate the sequence. The result is obtained by a decomposition of the final basis matrix.

1. Introduction

Pivoting methods for linear programming are based on solving a sequence of linear system of equations determined by a square submatrix of the constraint matrix, that typically changes by one column and/or one row in between iterations. The simplex method [2] is probably by far the most well-known, but we also want to mention criss-cross methods [1, 4, 7] and Lemke's method [6].

We consider a primal-dual pair of linear programs in canonical form

$$\begin{array}{ll}
 \text{minimize} & c^T x \\
 \text{subject to} & Ax \geq b, \\
 & x \geq 0,
 \end{array}
 \quad (PLP)
 \qquad
 \begin{array}{ll}
 \text{maximize} & b^T y \\
 \text{subject to} & A^T y \leq c, \\
 & y \geq 0.
 \end{array}
 \quad (DLP)$$

We show that if (PLP) and (DLP) are both feasible, then there exists a nonnegative integer r , with $r \leq \min\{m, n\}$, and a sequence of pivots (i_k, j_k) , $k = 1, \dots, r$, which generate sets of row indices $R_k = \cup_{l=1}^k \{i_l\}$ and columns indices $C_k = \cup_{l=1}^k \{j_l\}$, with

*Optimization and Systems Theory, Department of Mathematics, KTH Royal Institute of Technology, SE-100 44 Stockholm, Sweden (andersf@kth.se). Research supported by the Swedish Research Council (VR).

†Optimization and Systems Theory, Department of Mathematics, KTH Royal Institute of Technology, SE-100 44 Stockholm, Sweden (fewa@kth.se).

$i_{k+1} \in \{1, \dots, m\} \setminus R_k$ and $j_{k+1} \in \{1, \dots, n\} \setminus C_k$, such that $A_{R_k C_k}$ is nonsingular, and if x_{C_k} and y_{R_k} are computed from

$$A_{R_k C_k} x_{C_k} = b_{R_k}, \quad A_{R_k C_k}^T y_{R_k} = c_{C_k},$$

they are nonnegative, and therefore optimal to (PLP_k) and (DLP_k) respectively, where

$$(PLP_k) \quad \begin{array}{ll} \text{minimize} & c_{C_k}^T x_{C_k} \\ \text{subject to} & A_{R_k C_k} x_{C_k} \geq b_{R_k}, \\ & x_{C_k} \geq 0, \end{array} \quad (DLP_k) \quad \begin{array}{ll} \text{maximize} & b_{R_k}^T y_{R_k} \\ \text{subject to} & A_{R_k C_k}^T y_{R_k} \leq c_{C_k}, \\ & y_{C_k} \geq 0. \end{array}$$

Finally, x_{C_r} and y_{R_r} are not only optimal to (PLP_r) and (DLP_r) , but together with $x_j = 0$ for $j \in \{1, \dots, n\} \setminus C_r$ and $y_i = 0$ for $i \in \{1, \dots, m\} \setminus R_r$ they also give optimal solutions to (PLP) and (DLP) respectively. We refer to such a sequence of pivots as a *short* sequence of pivots. The existence of this short sequence of pivots is shown by a decomposition of the optimal basis matrix.

We also give a related result for a slightly more structured linear program, which is a min-max problem for a given $m \times n$ matrix M , formulated as the following primal-dual pair of linear programs

$$(P) \quad \begin{array}{ll} \text{minimize} & \alpha \\ \text{subject to} & Mu + e\alpha \geq 0, \\ & e^T u = 1, \\ & u \geq 0, \end{array} \quad (D) \quad \begin{array}{ll} \text{maximize} & \beta \\ \text{subject to} & M^T v + e\beta \leq 0, \\ & e^T v = 1, \\ & v \geq 0. \end{array}$$

For the general linear program, we cannot relate the short sequence of pivots to monotonicity in objective function value, whereas this can be done for the min-max problem. The difference is that (P) and (D) are both defined on the unit simplex, and they are both always feasible.

The results are straightforward, but to the best of our knowledge, our result on the existence of a sequence of pivots of length at most $\min\{m, n\}$ improves on what is known. Fukuda, Lühti and Namiki [3] and Fukuda and Terlaky [5] show that there is a sequence of pivots of length bounded by at most $m + n$ leading to the optimal solution. It should be pointed out that our existence result does not automatically give a method with better worst-case complexity than enumeration of all potential basis matrices. We have not been able to use the information to give rules based on global information that makes use of the short sequence of pivots.

2. Existence of a short pivoting sequence for a linear program

We will refer to a nonsingular square submatrix of A as a *basis matrix*. If R_+ and C_+ denote the row and column indices of A that define the basis matrix, the basis matrix is referred to as $A_{R_+ C_+}$. If R_0 and C_0 denote the remaining row and column indices, the primal and dual *basic solutions* associated with $A_{R_+ C_+}$ are uniquely

given by

$$\begin{aligned} A_{R_+C_+}x_{C_+} &= b_{R_+}, & A_{R_+C_+}^T y_{R_+} &= c_{C_+}, \\ x_{C_0} &= 0, & y_{R_0} &= 0. \end{aligned}$$

The primal-dual pair of basic solutions given by the basis matrix is optimal to (PLP) and (DLP) respectively if and only if the solutions are feasible to (PLP) and (DLP) respectively, i.e.,

$$\begin{aligned} A_{R_0C_+}x_{C_+} &\geq b_{R_0}, & A_{R_+C_0}^T y_{R_+} &\leq c_{C_0}, \\ x_{C_+} &\geq 0, & y_{R_+} &\geq 0. \end{aligned}$$

If (PLP) and (DLP) are both feasible, then there exists at least one basis matrix which gives a primal-dual optimal pair of basic solutions. This well-known result is summarized in the following lemma.

Lemma 2.1. (Existence of optimal basic feasible solution) *Assume that both (PLP) and (DLP) are feasible. Then, there is a partitioning of the row indices of A into two sets R_+ and R_0 , and a partitioning of the column indices of A into two sets C_+ and C_0 , so that $|R_+| = |C_+|$, and associated with the resulting matrix*

$$\begin{pmatrix} A_{R_+C_+} & A_{R_+C_0} \\ A_{R_0C_+} & A_{R_0C_0} \end{pmatrix},$$

the submatrix $A_{R_+C_+}$ is nonsingular, and there are vectors x and y for which

$$\begin{aligned} A_{R_+C_+}x_{C_+} &= b_{R_+}, & A_{R_+C_+}^T y_{R_+} &= c_{C_+}, \\ A_{R_0C_+}x_{C_+} &\geq b_{R_0}, & A_{R_+C_0}^T y_{R_+} &\leq c_{C_0}, \\ x_{C_+} &\geq 0, & y_{R_+} &\geq 0, \\ x_{C_0} &= 0, & y_{R_0} &= 0, \end{aligned}$$

hold. These vectors x and y are optimal solutions to (PLP) and (DLP) respectively.

Proof. Proofs are typically given for standard form of a linear program, e.g., [8, Theorem 3.4]. This can be achieved by adding slack variables to (PLP) , from which the result follows. ■

Our concern is to decompose the basis matrix by eliminating one row and one column at a time. The following lemma gives the basis for such an elimination of row i and column j for a given set of row indices R and column indices C .

Lemma 2.2. *Consider problems (PLP) and (DLP) . Let R denote a set of row indices of A and let C denote a set of column indices of A . Assume that*

$$A_{RC}x_C = b_R, \quad A_{RC}^T y_R = c_C, \quad A_{RC}\Delta x_C = e_i, \quad A_{RC}^T \Delta y_R = e_j, \quad (2.1)$$

where e_i and e_j are the i th and j th unit vectors of dimensions $|R|$ and $|C|$ respectively. Then,

$$c_C^T x_C = b_R^T y_R, \quad (2.2a)$$

$$c_C^T \Delta x_C = e_i^T y_R, \quad (2.2b)$$

$$b_R^T \Delta y_R = e_j^T x_C, \quad (2.2c)$$

$$e_j^T \Delta x_C = e_i^T \Delta y_R. \quad (2.2d)$$

In addition, assume that $e_j^T \Delta x_C = e_i^T \Delta y_R \neq 0$. Then there are unique scalars s_i and t_j such that

$$e_j^T (x_C + s_i \Delta x_C) = 0, \quad e_i^T (y_R + t_j \Delta y_R) = 0, \quad (2.3)$$

given by

$$s_i = -\frac{e_j^T x_C}{e_j^T \Delta x_C}, \quad t_j = -\frac{e_i^T y_R}{e_i^T \Delta y_R}. \quad (2.4)$$

Furthermore,

$$c_C^T (x_C + s_i \Delta x_C) = b_R^T (y_R + t_j \Delta y_R). \quad (2.5)$$

Proof. We obtain

$$\begin{aligned} c_C^T x_C &= y_R^T A_{RC} x_C = y_R^T b_R = b_R^T y_R, \\ c_C^T \Delta x_C &= y_R^T A_{RC} \Delta x_C = y_R^T e_i = e_i^T y_R, \\ b_R^T \Delta y_R &= x_C^T A_{RC}^T \Delta y_R = x_C^T e_j = e_j^T x_C, \\ e_j^T \Delta x_C &= \Delta y_R^T A_{RC} \Delta x_C = \Delta y_R^T e_i = e_i^T \Delta y_R. \end{aligned}$$

To show the final results, if $e_j^T \Delta x_C = e_i^T \Delta y_R \neq 0$, then the values of s_i and t_j given by (2.4) follow immediately. From these values of s_i and t_j , we obtain

$$\begin{aligned} 0 &= e_j^T (x_C + s_i \Delta x_C) = \Delta y_R^T A_{RC} (x_C + s_i \Delta x_C) = b_R^T \Delta y_R + s_i \Delta y_R^T A_{RC} \Delta x_C, \\ 0 &= e_i^T (y_R + t_j \Delta y_R) = \Delta x_C^T A_{RC}^T (y_R + t_j \Delta y_R) = c_C^T \Delta x_C + t_j \Delta x_C^T A_{RC}^T \Delta y_R. \end{aligned}$$

A combination of these equations gives

$$t_j b_R^T \Delta y_R = s_i c_C^T \Delta x_C = -s_i t_j \Delta x_C^T A_{RC}^T \Delta y_R. \quad (2.6)$$

A combination of $b_R^T y_R = c_C^T x_C$ and (2.6) gives (2.5), as required. ■

This result may now be used to reduce the dimension of the basis matrix by one row and one column, while maintaining primal and dual optimality to the reduced problem.

Lemma 2.3. Consider problems (PLP) and (DLP). Let R_k denote a set of row indices of A and let C_k denote a set of column indices of A such that $|R_k| = |C_k| = k$, with $k \geq 2$. Assume that $A_{R_k C_k}$ is nonsingular, and assume that

$$A_{R_k C_k} x_{C_k} = b_{C_k}, \quad A_{R_k C_k}^T y_{R_k} = c_{C_k},$$

where $x_{C_k} \geq 0$ and $y_{R_k} \geq 0$. Then, there is a row index i_k , with $i_k \in R_k$, and a column index j_k , with $j_k \in C_k$, such that $A_{R_{k-1}C_{k-1}}$ is nonsingular, where $R_{k-1} = R_k \setminus \{i_k\}$ and $C_{k-1} = C_k \setminus \{j_k\}$. Furthermore, it holds that

$$A_{R_{k-1}C_{k-1}}x_{C_{k-1}} = b_{C_{k-1}}, \quad A_{R_{k-1}C_{k-1}}^T y_{R_{k-1}} = c_{C_{k-1}},$$

for $x_{C_{k-1}} \geq 0$, $y_{R_{k-1}} \geq 0$.

Proof. We may apply Lemma 2.2 for $R = R_k$ and $C = C_k$. The quantities x_{C_k} , y_{R_k} , Δx_{C_k} and Δy_{R_k} are well defined since $A_{R_k C_k}$ is nonsingular.

First, assume that $y_i = 0$ for some $i \in R_k$. Let $i_k = i$. Compute Δx_{C_k} as in Lemma 2.2 for this i . If $\Delta x_{C_k} \not\leq 0$, we may compute the most limiting positive step for maintaining nonnegativity of $x_{C_k} + s\Delta x_{C_k}$, i.e.,

$$s = \min_{j \in C_k: e_j^T \Delta x_{C_k} < 0} \frac{e_j^T x_{C_k}}{-e_j^T \Delta x_{C_k}}. \quad (2.7)$$

If $\Delta x_{C_k} \not\leq 0$, we may compute the most limiting negative step for maintaining nonnegativity of $x_{C_k} + s\Delta x_{C_k}$, i.e.,

$$s = \max_{j \in C_k: e_j^T \Delta x_{C_k} > 0} \frac{e_j^T x_{C_k}}{-e_j^T \Delta x_{C_k}}. \quad (2.8)$$

Since $\Delta x_{C_k} \neq 0$, at least one of (2.7) and (2.8) is well defined. Pick one which is well defined, let j_k be a minimizing index and let s_i be the corresponding s -value. If $R_{k-1} = R_k \setminus \{i_k\}$ and $C_{k-1} = C_k \setminus \{j_k\}$, then $A_{R_{k-1}C_{k-1}}$ is nonsingular, since Δx_{C_k} up to a scalar is the unique vector in the nullspace of $A_{R_{k-1}C_k}$, and $e_{j_k}^T \Delta x_{C_k} \neq 0$ by (2.7) and (2.8).

Now assume that $x_j = 0$ for some $j \in C_k$. This is totally analogous to the case $y_i = 0$. Let $j_k = j$. Compute Δy_{R_k} as in Lemma 2.2 for this j . If $\Delta y_{R_k} \not\geq 0$, we may compute the most limiting positive step for maintaining nonnegativity of $y_{R_k} + t\Delta y_{R_k}$, i.e.,

$$t = \min_{i \in R_k: e_i^T \Delta y_{R_k} < 0} \frac{e_i^T y_{R_k}}{-e_i^T \Delta y_{R_k}}. \quad (2.9)$$

If $\Delta y_{R_k} \not\geq 0$, we may compute the most limiting negative step for maintaining nonnegativity of $y_{R_k} + t\Delta y_{R_k}$, i.e.,

$$t = \max_{i \in R_k: e_i^T \Delta y_{R_k} > 0} \frac{e_i^T y_{R_k}}{-e_i^T \Delta y_{R_k}}. \quad (2.10)$$

Since $\Delta y_{R_k} \neq 0$, at least one of (2.9) and (2.10) is well defined. Pick one which is well defined, let i_k be a minimizing index and let t_j be the corresponding t -value. If $R_{k-1} = R_k \setminus \{i_k\}$ and $C_{k-1} = C_k \setminus \{j_k\}$, then $A_{R_{k-1}C_{k-1}}$ is nonsingular, since Δy_{R_k} up to a scalar is the unique vector in the nullspace of $A_{R_k C_{k-1}}$, and $e_{i_k}^T \Delta y_{R_k} \neq 0$ by (2.9) and (2.10).

Finally, we consider the case when $x_{C_k} > 0$ and $y_{R_k} > 0$. Then, for any i, j pair with $i \in R_k$ and $j \in C_k$, Lemma 2.2 gives $c_{C_k}^T \Delta x_{C_k} > 0$ since $y_{R_k} > 0$ and $b_{R_k}^T \Delta y_{R_k} > 0$ since $x_{C_k} > 0$. We may now pick an $i \in R_k$ and compute Δx_{C_k} as in (2.1). We must have $\Delta x_{C_k} \neq 0$, since $e_i \neq 0$. We may now compute the most limiting step from (2.7) or (2.8), out of which at least one has to be well defined. Assume first the former. Let s_i denote the step and let j be an index for which the minimum is attained, so that $e_j^T(x_{C_k} + s_i \Delta x_{C_k}) = 0$. By computing Δy_{R_k} for this j , there is an associated positive t_j such that $e_i^T(y_{R_k} + t_j \Delta y_{R_k}) = 0$ by Lemma 2.2. If $y_{R_k} + t_j \Delta y_{R_k} \geq 0$, we are done. Otherwise, we let t_j be the maximum positive step such that $y_{R_k} + t_j \Delta y_{R_k} \geq 0$. By construction, this must give a strict reduction of t_j , but t_j will remain strictly positive since $y_{R_k} > 0$. We now conversely find the associated step s_i . This process may be repeated a finite number of times for i, j pairs until $x_{C_k} + s_i \Delta x_{C_k} \geq 0$ with $e_j^T(x_{C_k} + s_i \Delta x_{C_k}) = 0$ and $y_{R_k} + t_j \Delta y_{R_k} \geq 0$ with $e_i^T(y_{R_k} + t_j \Delta y_{R_k}) = 0$. Note that (2.5) of Lemma 2.2 implies that each step gives a strict reduction in objective function value, since one of the s_i or t_j values is reduced. Let $i_k = i$ and $j_k = j$. If $R_{k-1} = R_k \setminus \{i_k\}$ and $C_{k-1} = C_k \setminus \{j_k\}$, then $A_{R_{k-1}C_{k-1}}$ is nonsingular, since Δx_{C_k} up to a scalar is the unique vector in the nullspace of $A_{R_{k-1}C_k}$, and $e_{j_k}^T \Delta x_{C_k} \neq 0$. If (2.10) is used instead of (2.9), the argument is analogous, but now s_i and t_j are negative, increasing towards zero. ■

The optimality conditions given by Lemma 2.1 imply the existence of a nonsingular submatrix $A_{R_+C_+}$. Recursive application of Lemma 2.3 gives a decomposition of this matrix into square nonsingular submatrices of dimensions shrinking by one row and one column at a time, corresponding to primal-dual optimal pairs of (PLP_k) and (DLP_k) respectively, for $k = r, r-1, \dots, 1$. By reversing this argument, there must exist a sequence of r pivots $(i_1, j_1), (i_2, j_2), \dots, (i_r, j_r)$, such that at stage k , optimal solutions to (PLP_k) and (DLP_k) are created, and at stage r , optimal solutions to (PLP) and (DLP) are found. This is summarized in the following theorem.

Theorem 2.1. *Assume that problems (PLP) and (DLP) both have feasible solutions. For the optimality conditions given by Lemma 2.1, let $r = |R_+| = |C_+|$. Then, $r \leq \min\{m, n\}$ and there are pairs of row and column indices (i_k, j_k) , $k = 1, \dots, r$, which generate sets of row indices $R_1 = \{i_1\}$, $R_{k+1} = R_k \cup \{i_{k+1}\}$, and sets of column indices $C_1 = \{j_1\}$, $C_{k+1} = C_k \cup \{j_{k+1}\}$, with $i_{k+1} \in \{1, \dots, m\} \setminus R_k$ and $j_{k+1} \in \{1, \dots, n\} \setminus C_k$, such that for each k , $A_{R_k C_k}$ is nonsingular, and x_{C_k} and y_{R_k} computed from*

$$A_{R_k C_k} x_{C_k} = b_{R_k}, \quad A_{R_k C_k}^T y_{R_k} = c_{C_k},$$

are optimal to (PLP_k) and (DLP_k) respectively. In addition, x_{C_r} and y_{R_r} together with $x_j = 0$ for $j \in \{1, \dots, n\} \setminus C_r$ and $y_i = 0$ for $i \in \{1, \dots, m\} \setminus R_r$ are optimal to (PLP) and (DLP) respectively.

Proof. For $r = 0$ or $r = 1$, the result is immediate from the optimality conditions of Lemma 2.1. For $r \geq 2$, let $R_r = R_+$ and $C_r = C_+$, and repeatedly apply Lemma 2.3 for $k = r, r-1, r-2, \dots, 2$. This gives an ordering of the indices of R_+ and C_+ as

i_r, i_{r-1}, \dots, i_1 and j_r, j_{r-1}, \dots, j_1 , such that the corresponding x_{C_k} and y_{R_k} are optimal to (PLP_k) and (DLP_k) respectively for $k = 1, \dots, r$. In addition, x_{C_r} and y_{R_r} are optimal to (PLP) and (DLP) respectively. If the ordering is reversed, so that $k = 1, \dots, r$, the result follows. ■

We note that Theorem 2.1 shows the existence of a sequence of pivots that would create the optimal basis in r steps, where $r \leq \min\{m, n\}$. We refer to such a sequence of pivots as a *short* sequence of pivots. This, however, does not constitute an algorithm based on global information. We have no global information on how to create the sequence of pivots. There is a straightforward method given by enumerating all potential sequences of pivots that generate primal-dual optimal pairs of linear programs of increasing dimension. However, we have not been able to give any useful bound on the potential number of bases.

We also note that the short sequence of pivots does not explicitly take into account objective function value, as the simplex method does. Therefore, we cannot ensure generating a short sequence of pivots by making use of pivots associated with primal simplex only or dual simplex only.

3. Decomposing a matrix by convex combinations

We also consider the related problem of decomposing a given $m \times n$ matrix M by convex combination, formulated as the following primal-dual pair of linear programs

$$(P) \quad \begin{array}{ll} \underset{u \in \mathbb{R}^n, \alpha \in \mathbb{R}}{\text{minimize}} & \alpha \\ \text{subject to} & Mu + e\alpha \geq 0, \\ & e^T u = 1, \\ & u \geq 0, \end{array} \quad (D) \quad \begin{array}{ll} \underset{v \in \mathbb{R}^m, \beta \in \mathbb{R}}{\text{maximize}} & \beta \\ \text{subject to} & M^T v + e\beta \leq 0, \\ & e^T v = 1, \\ & v \geq 0. \end{array}$$

Here, and throughout, e denotes the vector of ones of the appropriate dimension.

Problems (P) and (D) have a joint optimal value γ by strong duality for linear programming. The difference to the general linear program is that u and v are defined on the unit simplex, and (P) and (D) are always feasible. This will enable a slightly stronger result in which monotonicity in the objective function value may be enforced.

We state the analogous results to the general linear programming case, and point out the differences.

Lemma 3.1. (Existence of optimal basic feasible solution) *For a given matrix M , there exists a number γ and a partitioning of the row indices of M into two sets R_+ and R_0 , and a partitioning of the column indices of M into two sets C_+ and C_0 , so that $|R_+| = |C_+|$, and associated with the resulting matrix*

$$\begin{pmatrix} M_{R_+C_+} & M_{R_+C_0} & e \\ M_{R_0C_+} & M_{R_0C_0} & e \\ e^T & e^T & 0 \end{pmatrix},$$

the submatrix

$$\begin{pmatrix} M_{R_+C_+} & e \\ e^T & 0 \end{pmatrix}$$

is nonsingular, and there are vectors u and v for which

$$\begin{aligned} \begin{pmatrix} M_{R_+C_+} & e \\ e^T & 0 \end{pmatrix} \begin{pmatrix} u_{C_+} \\ \gamma \end{pmatrix} &= \begin{pmatrix} 0 \\ 1 \end{pmatrix}, & \begin{pmatrix} M_{R_+C_+}^T & e \\ e^T & 0 \end{pmatrix} \begin{pmatrix} v_{R_+} \\ \gamma \end{pmatrix} &= \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \\ \begin{pmatrix} M_{R_0C_+} & e \end{pmatrix} \begin{pmatrix} u_{C_+} \\ \gamma \end{pmatrix} &\geq 0, & \begin{pmatrix} M_{R_+C_0}^T & e \end{pmatrix} \begin{pmatrix} v_{R_+} \\ \gamma \end{pmatrix} &\leq 0, \\ u_{C_+} &\geq 0, & v_{R_+} &\geq 0, \\ u_{C_0} &= 0, & v_{R_0} &= 0, \end{aligned}$$

hold. The vectors (u, γ) and (v, γ) are optimal solutions to (P) and (D) respectively.

Proof. This is analogous to Lemma 2.1. The only difference is that α and β are free variables. ■

Lemma 3.2. Consider problems (P) and (D). Let R denote a set of row indices of M and let C denote a set of column indices of M . Assume that

$$\begin{aligned} \begin{pmatrix} M_{RC} & e \\ e^T & 0 \end{pmatrix} \begin{pmatrix} u_C \\ \alpha \end{pmatrix} &= \begin{pmatrix} 0 \\ 1 \end{pmatrix}, & \begin{pmatrix} M_{RC}^T & e \\ e^T & 0 \end{pmatrix} \begin{pmatrix} v_R \\ \beta \end{pmatrix} &= \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \\ \begin{pmatrix} M_{RC} & e \\ e^T & 0 \end{pmatrix} \begin{pmatrix} \Delta u_C \\ \Delta \alpha \end{pmatrix} &= \begin{pmatrix} e_i \\ 0 \end{pmatrix}, & \begin{pmatrix} M_{RC}^T & e \\ e^T & 0 \end{pmatrix} \begin{pmatrix} \Delta v_R \\ \Delta \beta \end{pmatrix} &= \begin{pmatrix} e_j \\ 0 \end{pmatrix}. \end{aligned}$$

Then,

$$\alpha = \beta, \tag{3.1a}$$

$$\Delta \alpha = e_i^T v_R, \tag{3.1b}$$

$$\Delta \beta = e_j^T u_C, \tag{3.1c}$$

$$e_j^T \Delta u_C = e_i^T \Delta v_R. \tag{3.1d}$$

In addition, assume that $e_j^T \Delta u_C = e_i^T \Delta v_R \neq 0$. Then there are unique scalars s_i and t_j such that

$$e_j^T (u_C + s_i \Delta u_C) = 0, \quad e_i^T (v_R + t_j \Delta v_R) = 0,$$

given by

$$s_i = -\frac{e_j^T u_C}{e_j^T \Delta u_C}, \quad t_j = -\frac{e_i^T v_R}{e_i^T \Delta v_R}.$$

Furthermore,

$$\alpha + s_i \Delta \alpha = \beta + t_j \Delta \beta.$$

Proof. This is analogous to Lemma 2.2. ■

As for the general linear programming case, associated with (P) and (D) we will consider the linear programs

$$(P_k) \quad \begin{array}{ll} \text{minimize} & \alpha_k \\ u_{C_k} \in \mathbb{R}^k, \alpha_k \in \mathbb{R} & \\ \text{subject to} & M_{R_k C_k} u_{C_k} + e \alpha_k \geq 0, \\ & e^T u_{C_k} = 1, \\ & u_{C_k} \geq 0, \end{array} \quad (D_k) \quad \begin{array}{ll} \text{maximize} & \beta_k \\ v_{R_k} \in \mathbb{R}^k, \beta_k \in \mathbb{R} & \\ \text{subject to} & M_{R_k C_k}^T v_{C_k} + e \beta_k \leq 0, \\ & e^T v_{C_k} = 1, \\ & v_{C_k} \geq 0. \end{array}$$

where R_k denotes a set of row indices of A and C_k denotes a set of column indices of M such that $|R_k| = |C_k| = k$.

Lemma 3.3. *Let M be an $m \times n$ matrix. Let R_k denote a set of row indices of M and let C_k denote a set of column indices of M such that $|R_k| = |C_k| = k$, with $k \geq 2$. Assume that*

$$\begin{pmatrix} M_{R_k C_k} & e \\ e^T & 0 \end{pmatrix}$$

is nonsingular, and assume that

$$\begin{pmatrix} M_{R_k C_k} & e \\ e^T & 0 \end{pmatrix} \begin{pmatrix} u_{C_k} \\ \gamma_k \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad \begin{pmatrix} M_{R_k C_k}^T & e \\ e^T & 0 \end{pmatrix} \begin{pmatrix} v_{R_k} \\ \gamma_k \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix},$$

where $u_{C_k} \geq 0$ and $v_{R_k} \geq 0$.

Then, there is a row index $i_k^{(1)}, i_k^{(1)} \in R_k$, and a column index $j_k^{(1)}, j_k^{(1)} \in C_k$, such that if $R_{k-1}^{(1)} = R_k \setminus \{i_k^{(1)}\}$ and $C_{k-1}^{(1)} = C_k \setminus \{j_k^{(1)}\}$, then

$$\begin{pmatrix} M_{R_{k-1}^{(1)} C_{k-1}^{(1)}} & e \\ e^T & 0 \end{pmatrix}$$

is nonsingular. Furthermore, it holds that

$$\begin{pmatrix} M_{R_{k-1}^{(1)} C_{k-1}^{(1)}} & e \\ e^T & 0 \end{pmatrix} \begin{pmatrix} u_{C_{k-1}^{(1)}} \\ \gamma_{k-1}^{(1)} \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad \begin{pmatrix} (M_{R_{k-1}^{(1)} C_{k-1}^{(1)}})^T & e \\ e^T & 0 \end{pmatrix} \begin{pmatrix} v_{R_{k-1}^{(1)}} \\ \gamma_{k-1}^{(1)} \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix},$$

for $u_{C_{k-1}^{(1)}} \geq 0$, $v_{R_{k-1}^{(1)}} \geq 0$ and $\gamma_{k-1}^{(1)} \leq \gamma_k^{(1)}$.

In addition, there is a row index $i_k^{(2)}, i_k^{(2)} \in R_k$, and a column index $j_k^{(2)}, j_k^{(2)} \in C_k$, with $(i_k^{(2)}, j_k^{(2)}) \neq (i_k^{(1)}, j_k^{(1)})$, such that if $R_{k-1}^{(2)} = R_k \setminus \{i_k^{(2)}\}$ and $C_{k-1}^{(2)} = C_k \setminus \{j_k^{(2)}\}$, then

$$\begin{pmatrix} M_{R_{k-1}^{(2)} C_{k-1}^{(2)}} & e \\ e^T & 0 \end{pmatrix}$$

is nonsingular. Furthermore, it holds that

$$\begin{pmatrix} M_{R_{k-1}C_{k-1}}^{(2)} & e \\ e^T & 0 \end{pmatrix} \begin{pmatrix} u_{C_{k-1}}^{(2)} \\ \gamma_{k-1}^{(2)} \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad \begin{pmatrix} (M_{R_{k-1}C_{k-1}}^{(2)})^T & e \\ e^T & 0 \end{pmatrix} \begin{pmatrix} v_{R_{k-1}}^{(2)} \\ \gamma_{k-1}^{(2)} \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix},$$

for $u_{C_{k-1}}^{(2)} \geq 0$, $v_{R_{k-1}}^{(2)} \geq 0$ and $\gamma_{k-1}^{(2)} \geq \gamma_k^{(2)}$.

Proof. The difference compared to Lemma 2.3 is that α_k and β_k are free variables. Hence, there is no nonnegativity condition on them to handle. We note from Lemma 3.2 that

$$\begin{pmatrix} M_{R_k C_k} & e \\ e^T & 0 \end{pmatrix} \begin{pmatrix} \Delta u_{C_k} \\ \Delta \alpha_k \end{pmatrix} = \begin{pmatrix} e_i \\ 0 \end{pmatrix}.$$

Hence, we cannot have $\Delta u_{C_k} = 0$, since $e \Delta \alpha_k = e_i$ cannot have a solution for $k \geq 2$. It follows that $\Delta u_{C_k} \neq 0$ so that $e^T \Delta u_{C_k} = 0$ implies that Δu_{C_k} must have both strictly positive and strictly negative components. The situation is analogous for Δv_{R_k} . Therefore, there is a choice of selecting s negative or positive analogously to (2.7) and (2.8). Consequently, there are two different possible index pairs, one corresponding to $\Delta \gamma_k \geq 0$ and one corresponding to $\Delta \gamma_k \leq 0$. ■

Theorem 3.1. *Let M be a given $m \times n$ matrix. For the optimality conditions given by Lemma 3.1, let $r = |R_+| = |C_+|$. Then, $r \leq \min\{m, n\}$ and there are pairs of row and column indices (i_k, j_k) , $k = 1, \dots, r$, which generate sets of row indices $R_1 = \{i_1\}$, $R_{k+1} = R_k \cup \{i_{k+1}\}$, and sets of column indices $C_1 = \{j_1\}$, $C_{k+1} = C_k \cup \{j_{k+1}\}$, with $i_{k+1} \in \{1, \dots, m\} \setminus R_k$ and $j_{k+1} \in \{1, \dots, n\} \setminus C_k$, such that for each k ,*

$$\begin{pmatrix} M_{R_k C_k} & e \\ e^T & 0 \end{pmatrix}$$

is nonsingular, and (u_{C_k}, γ_k) and (v_{R_k}, γ_k) computed from

$$\begin{pmatrix} M_{R_k C_k} & e \\ e^T & 0 \end{pmatrix} \begin{pmatrix} u_{C_k} \\ \gamma_k \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad \begin{pmatrix} M_{R_k C_k}^T & e \\ e^T & 0 \end{pmatrix} \begin{pmatrix} v_{R_k} \\ \gamma_k \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix},$$

are optimal to (P_k) and (D_k) respectively. In addition, (u_{C_r}, γ_r) and (v_{R_r}, γ_r) together with $u_j = 0$ for $j \in \{1, \dots, n\} \setminus C_r$ and $v_i = 0$ for $i \in \{1, \dots, m\} \setminus R_r$ are optimal to (P) and (D) respectively.

Such a sequence of pairs of row and column indices (i_k, j_k) , $k = 1, \dots, r$, exists also if one of the additional requirements $\gamma_{k+1} \leq \gamma_k$, $k = 1, \dots, r-1$, or $\gamma_{k+1} \geq \gamma_k$, $k = 1, \dots, r-1$, are imposed.

Proof. The result is analogous to Theorem 2.1. The only difference is the final statement on additional requirement $\gamma_{k+1} \leq \gamma_k$, $k = 1, \dots, r-1$, or $\gamma_{k+1} \geq \gamma_k$, $k = 1, \dots, r-1$, not contradicting the existence of the short pivot sequence. This is a consequence of the proved existence of two potential pivots in the reduction step of Lemma 3.3. ■

4. Summary

For a pair of linear programs in canonical forms that both are feasible, we have shown the existence of a sequence of pivots of length at most $\min\{m, n\}$ that leads from the origin to a primal-dual pair of optimal solutions. At each step, the pivot creates a nonsingular submatrix of the constraint matrix that increases in dimension by one row and column by including the row and column of the pivot element. By solving two linear systems involving the submatrix a pair of primal and dual solutions are obtained. These solutions are optimal for the restricted problem where only rows and columns of the submatrix are included. At the final step, the solutions are optimal to the full primal and dual problem respectively.

We have not been able to give rules for an algorithm taking into account global information that would give this correct path without potentially enumerating all possible paths, which might be exponentially many. We therefore only publish the result as is, and hope that the result will be useful for further understanding of pivoting methods for linear programming.

We also note in passing that the reduction of Lemma 2.3 and Lemma 3.3 can be done from an arbitrary basis matrix if the nonnegativity condition is omitted. Therefore, there is a sequence of pivots of length at most $\min\{m, n\}$ leading from any pair of primal-dual basic solutions to the origin, Consequently, Theorem 2.1 and Theorem 3.1 imply that there is an overall bound of at most $2 \min\{m, n\}$ pivots leading from any pair of primal-dual basic solutions to an optimal pair of primal-dual basic solutions. This is at least as tight as the bound $n + m$ given by Fukuda and Terlaky [5].

Acknowledgement

We thank Krister Svanberg for many helpful discussions.

References

- [1] T. Bonates and N. Maculan. Performance evaluation of a family of criss-cross algorithms for linear programming. *Int. Trans. Oper. Res.*, 10:53–64, 2003.
- [2] G. B. Dantzig. *Linear Programming and Extensions*. Princeton University Press, Princeton, New Jersey, 1963.
- [3] K. Fukuda, H.-J. Lüthi, and M. Namiki. The existence of a short sequence of admissible pivots to an optimal basis in LP and LCP. *Int. Trans. Oper. Res.*, 4:273–284, 1997.
- [4] K. Fukuda and T. Terlaky. Criss-cross methods: A fresh view on pivot algorithms. *Math. Prog.*, 79:369–395, 1997.
- [5] K. Fukuda and T. Terlaky. On the existence of a short admissible pivot sequence for feasibility and linear optimization problems. *Pure Math. Appl.*, 10:431–447, 1999.
- [6] C. E. Lemke. The dual method of solving the linear programming problem. *Naval Research Logistics Quarterly*, 1(1):36–47, 1954.
- [7] T. Terlaky. A convergent criss-cross method. *Optimization*, 16:683–690, 1985.
- [8] R. J. Vanderbei. *Linear programming*, volume 196 of *International Series in Operations Research & Management Science*. Springer, New York, fourth edition, 2014. Foundations and extensions.