AN ELEMENTARY PROOF OF OPTIMALITY CONDITIONS FOR LINEAR PROGRAMMING

Anders FORSGREN*

Technical Report TRITA-MAT-2008-OS6
Department of Mathematics
Royal Institute of Technology
June 2008

Abstract

We give an elementary proof of optimality conditions for linear programming. The proof is direct, built on a straightforward classical perturbation of the constraints, and does not require either the use of Farkas' lemma or the use of the simplex method.

1. Introduction

In this note, optimality conditions are derived for the linear programming problem (LP) defined in standard form as

$$(LP) \qquad \begin{array}{ll} \displaystyle\operatorname*{minimize}_{x \in I\!\!R^n} & c^T x \\ \text{subject to} & Ax = b, \\ & x \geq 0, \end{array}$$

where A is an $m \times n$ matrix, b is an m-dimensional vector and c is an n-dimensional vector. For convenience, it is assumed throughout that A has full row rank. A linear program may be posed on many different forms. The particular form is not important, since different forms may be transformed into each other. To make the exposition simple in this note, standard form is used throughout.

Traditionally, textbooks on linear programming often give optimality conditions for a linear program derived via Farkas' lemma, e.g., Shriver [10] and Gill, Murray and Wright [5], or via the simplex method, taking into account anti-cycling, e.g., Chvátal [2], Luenberger [8] and Nash and Sofer [9]. Linear programming and the simplex method have been very close together since the introduction of the simplex method in the 1940's by Dantzig, which is also reflected in his classical textbook on linear programming [3]. In particular, linear programming and the simplex method were almost interchangeable terms prior to the ellipsoid method of Khachian [7] in 1979. The method proposed by Karmarkar [6] in 1984, and subsequent development of interior methods for linear programming, has given a new situation where the

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

simplex method has competitors. The purpose of this note is to give an elementary proof of optimality conditions for linear programming, that does not need either Farkas' lemma or the simplex method. It has been inspired by the paper of Dax [4] and the manuscript of Svanberg [11], which give elementary proofs of Farkas' lemma.

A feature of linear programming is that if there is an optimal solution, there is an optimal *extreme point* of the polytope that defines the feasible region. An extreme point is a feasible point to (LP) which is uniquely defined by the active constraints. This means that at least n constraints have to be active. An extreme point is said to be *nondegenerate* if exactly n constraints are active. The problem is referred to as nondegenerate if every extreme point is nondegenerate. For a linear program on standard form, an extreme point is often referred to as a *basic feasible solution.*

Under a nondegeneracy assumption, the derivation of optimality conditions is straightforward. Covering the degenerate case as well requires some additional mathematics, which often gives a detour. One may derive some additional result like Farkas' lemma or introduce a method for linear programming, like the simplex method, and then introduce *anticycling schemes* in the simplex method to handle nondegeneracy. Although these approaches certainly have their merit, our intention is to derive optimality conditions directly with basic tools only, and then leave out more advanced features. The results which we derive are by no means original, but we are not aware of a direct treatment along the lines presented here. For the degenerate case, we rely on a perturbed problem introduced by Charnes [1], but we do not need to introduce the simplex method or the anti-cycling procedure. It suffices to study the perturbed problem.

1.1. Notation and terminology

A feasible point x to (LP) will be partitioned as $x = (x_+^T \; x_0^T)^T$, where $x_+ > 0$ and $x_0 = 0$. The ordering of the indices in x_+ and x_0 is not important, the positive components are collected in x_+ and the zero components are collected in x_0. In addition, the columns of A and the vector c will be partitioned conformally, i.e., $A = (A_+ \; A_0)$ and $c = (c_+^T \; c_0^T)^T$. Following standard terminology for linear programming, we will refer to a point x that satisfies $Ax = b$ at which A_+ has full column rank as a *basic solution.* If, in addition, x is nonnegative, it will be referred to as a *basic feasible solution.* We will also refer to the basic feasible solutions as the extreme points of the feasible region. The terms basic feasible solution and extreme point will be used interchangeably.

Associated with a basic feasible solution, we may partition $A = (B \; N)$, where the *basis matrix* B is $m \times m$ and nonsingular such that A_+ is a submatrix of B. Again, the ordering of the columns is not important, B is assumed to be formed by the m first columns for convenience in notation. For a column index j, we will denote by A_j the jth column of A. We will also use the notation $j \in B$ and $j \in A_+$ to denote that A_j is a column of B and A_+ respectively.

If the basic feasible solution is *nondegenerate*, then B is unique with $A_+ = B$. Otherwise, any nonsingular $m \times m$ submatrix B of A such that A_+ is a submatrix of

B gives a basis matrix. The reason why optimality conditions are more complicated for the degenerate case is that not every basis matrix may be used to form the optimality conditions then.

2. Sufficient optimality conditions for a linear program

It is straightforward to derive sufficient optimality conditions for (LP). These are stated in the following proposition.

Proposition 2.1. *Assume that there are $y \in I\!R^m$, $s \in I\!R^n$ such that*

$$A^T y + s = c, \tag{2.1a}$$

$$s \geq 0. \tag{2.1b}$$

Then, if x is feasible to (LP) it holds that

$$c^T x - b^T y = s^T x \geq 0. \tag{2.2}$$

In particular, if $s^T x = 0$, then x is optimal to (LP) and the optimal value is given by $b^T y$.

Proof. Assume that y and s satisfy (2.1a) and (2.1b). Then, for any x that is feasible to (LP), it holds that

$$c^T x = (s + A^T y)^T x = s^T x + b^T y \geq b^T y, \tag{2.3}$$

where the last inequality follows since x and s both are nonnegative. Hence, (2.2) holds and $b^T y$ is a lower bound on the optimal value of (LP). If in addition $s^T x = 0$, then the lower bound $b^T y$ given by (2.3) is attained, and x is thus optimal to (LP) with optimal value given by $b^T y$. ∎

3. Existence of an optimal basic feasible solution

We first show that unless the optimal value of (LP) is unbounded from below, (LP) has at least one optimal solution which is a basic feasible solution.

Lemma 3.1. *Assume that (LP) is feasible. Then, either (LP) has at least one optimal basic feasible solution or there is a $p \in I\!R^n$ such that $p \geq 0$, $Ap = 0$, $c^T p < 0$.*

Proof. Let x be a feasible point. If the associated A_+ has full column rank, then x is a basic feasible solution, and hence there is a basic feasible solution x_e such that $c^T x_e \leq c^T x$. If A_+ does not have full column rank, there is a nonzero vector p_+ such that $A_+ p_+ = 0$. The sign of p_+ may be chosen such that $c_+^T p_+ \leq 0$. There are now two cases, (i) $c_+^T p_+ < 0$ or (ii) $c_+^T p_+ = 0$. Case (i) $c_+^T p_+ < 0$ can be split in two subcases, (ia) $c_+^T p_+ < 0$ and $p_+ \geq 0$ or (ib) $c_+^T p_+ < 0$ and $p_+ \not\geq 0$. The three cases (ia), (ib) and (ii) will now be considered in turn.

Case (ia): $c_+^T p_+ < 0$ and $p_+ \geq 0$. In this situation, let $p_0 = 0$. Then, the resulting p satisfies $p \geq 0$, $Ap = 0$ and $c^T p < 0$.

Case (ib): $c_+^T p_+ < 0$ and $p_+ \not\geq 0$. Let

$$\alpha = \min_{i:(p_+)_i < 0} \frac{(x_+)_i}{-(p_+)_i}. \tag{3.1}$$

Since $p_+ \not\geq 0$, α is well defined. Let $p_0 = 0$. Then, the resulting p gives a point $x + \alpha p$ such that $A(x + \alpha p) = b$, $x + \alpha p \geq 0$ and $c^T(x + \alpha p) < c^T x$. In addition, $x + \alpha p$ has at least one more component which is zero compared to x.

Case (ii): $c_+^T p_+ = 0$. Since $p_+ \neq 0$ the sign of p_+ may be chosen such that $p_+ \not\geq 0$. It will still hold that $c_+^T p_+ = 0$. As in case (ib), we may now define α from (3.1) and let $p_0 = 0$. Then, the resulting p gives a point $x + \alpha p$ such that $A(x + \alpha p) = b$, $x + \alpha p \geq 0$ and $c^T(x + \alpha p) = c^T x$. In addition, $x + \alpha p$ has at least one more component which is zero compared to x.

Hence, if case (ia) does not occur, a new feasible point is obtained at which the objective function is no greater than $c^T x$ and which has at least one more component which is zero compared to x. This procedure may now be repeated for the new points generated. After at most n steps, either case (ia) has been reached at some step, or it has been concluded that A_+ for the generated point has full column rank. In the former case, a vector $p \in I\!R^n$ such that $p \geq 0$, $Ap = 0$, $c^T p < 0$, has been found. In the latter case, we have found a basic feasible solution x_e such that $c^T x_e \leq c^T x$.

Consequently, we conclude that for every feasible x, either a p such that $p \geq 0$, $Ap = 0$, $c^T p < 0$ is found, or it holds that $c^T x \geq \min_{x_e \in X_e} c^T x_e$, where X_e denotes the (finite) set of basic feasible solutions. Hence, in the latter case, at least one basic feasible solution is optimal. ∎

4. Existence of an optimal basis

The existence of an optimal basic feasible solution may now be used to derive necessary optimality conditions. First, optimality conditions for a nondegenerate optimal basic feasible solution are derived. The nonegeneracy implies that the basis matrix may be used to characterize the feasible neighborhood of the solution.

Lemma 4.1. *Assume that* (LP) *has a nondegenerate optimal basic feasible solution, i.e., there is an optimal basic feasible solution* x *for which there exists a basis matrix* B *such that*

$$\begin{aligned} &B x_B = b, \quad x_N = 0, \\ &x_B > 0. \end{aligned}$$

Then there exist vectors $y \in I\!R^m$ *and* $s \in I\!R^n$*, which may be partitioned conformally with* x *such that*

$$B^T y = c_B, \quad s_B = 0, \quad N^T y + s_N = c_N, \tag{4.1a}$$

$$s_N \geq 0. \tag{4.1b}$$

Proof. Let $\mathcal{P} = \{p \in I\!\!R^n : Bp_B + Np_N = 0,\ p_N \geq 0\}$. Then, since $x_B > 0$, it follows that $p \in \mathcal{P}$ is a feasible direction, i.e., there is a positive $\bar{\alpha}$ such that $x + \alpha p$ is feasible to (LP) for $\alpha \in [0, \bar{\alpha}]$. Consequently, since x is optimal, it must hold that $c^T p \geq 0$ for $p \in \mathcal{P}$. Hence,

$$0 \leq c^T p = c_B^T p_B + c_N^T p_N = (c_N - N^T B^{-T} c_B)^T p_N. \tag{4.2}$$

For the given B, we may define y and s from (4.1a). It remains to show that (4.1b) holds. Insertion of (4.1a) into (4.2) gives

$$0 \leq c^T p = s_N^T p_N. \tag{4.3}$$

Since p_N is an arbitrary nonnegative vector, (4.3) implies that $s_N \geq 0$, so that (4.1b) holds. ∎

For the degenerate case, the above analysis characterizing the feasible neighborhood in terms of a basis matrix is not straightforward. This is exactly the reason why giving optimality conditions for linear programming is not entirely straightforward. We now extend the result above to show the existence of an optimal basic feasible solution to (LP) for which the sufficient optimality conditions of Proposition 2.1 hold also to the degenerate case. This is done by a classical perturbation of the feasible region, originally suggested by Charnes [1] in the more complex context of handling degeneracy in the simplex method. The essence is that the constraints are perturbed by different amounts, thereby removing the degeneracy.

Proposition 4.1. *Assume that (LP) is feasible, and that there is no $p \in I\!\!R^n$ such that $p \geq 0$, $Ap = 0$, $c^T p < 0$. Then, there exists an optimal basic feasible solution $x \in I\!\!R^n$ for which there is an associated basis matrix B and vectors $y \in I\!\!R^m$ and $s \in I\!\!R^n$, such that*

$$Bx_B = b, \quad x_N = 0, \tag{4.4a}$$
$$x_B \geq 0, \tag{4.4b}$$
$$B^T y = c_B, \quad s_B = 0, \quad N^T y + s_N = c_N, \tag{4.4c}$$
$$s_N \geq 0. \tag{4.4d}$$

Proof. Since there is no $p \in I\!\!R^n$ such that $p \geq 0$, $Ap = 0$, $c^T p < 0$, Lemma 3.1 ensures the existence of at least one optimal basic feasible solution to (LP). If there exists a nondegenerate optimal basic feasible solution, Lemma 4.1 directly gives the desired result.

If (LP) does not have a nondegenerate optimal basic feasible solution, the situation is a bit more complicated. This degeneracy may be resolved by creating a perturbed problem as

$$\begin{array}{ll} \underset{x \in I\!\!R^n}{\text{minimize}} & c^T x \\ \text{subject to} & Ax = b, \\ & x_j \geq -\epsilon^j, \quad j = 1, \dots, n. \end{array} \tag{4.5}$$

for ϵ sufficiently small and positive. Let

$$\gamma = \min_{B\in\mathcal{B}} \left\{ \min_{i=1,\dots,m} \min_{\substack{j\in N \\ e_i^T B^{-1} A_j \neq 0}} |e_i^T B^{-1} A_j|, \min_{\substack{i=1,\dots,m \\ e_i^T B^{-1} b \neq 0}} |e_i^T B^{-1} b|, 1 \right\},$$

$$\Gamma = \max_{B\in\mathcal{B}} \left\{ \max_{i=1,\dots,m} \max_{\substack{j\in N \\ e_i^T B^{-1} A_j \neq 0}} |e_i^T B^{-1} A_j|, \max_{\substack{i=1,\dots,m \\ e_i^T B^{-1} b \neq 0}} |e_i^T B^{-1} b|, 1 \right\},$$

where $\mathcal{B}$ is the set of nonsingular $m \times m$ submatrices of A, and let $0 < \epsilon < \gamma/(\gamma+\Gamma)$.

Since (LP) is feasible and $\epsilon > 0$, it follows that (4.5) is feasible. By a change of variables, $\widetilde{x}_j = x_j + \epsilon_j$, $j = 1, \dots, n$, (4.5) may be rewritten in standard form as

$$(LP_\epsilon) \qquad \begin{array}{ll} \underset{\widetilde{x}\in I\!R^n}{\text{minimize}} & c^T \widetilde{x} \\ \text{subject to} & A\widetilde{x} = b + \sum_{j=1}^n A_j \epsilon^j, \\ & \widetilde{x} \ge 0. \end{array}$$

The objective functions in (4.5) and (LP_ϵ) differ by a constant, which does not have any impact on the solution. In addition, since there is no p such that $p \ge 0$, $Ap = 0$, $c^T p < 0$, Lemma 3.1 shows that (LP_ϵ) has at least one optimal basic feasible solution. Let $\widetilde{x}$ be such an optimal basic feasible solution, and let B be a corresponding basis matrix, i.e., a nonsingular $m \times m$ submatrix of A such that the A_+ associated with $\widetilde{x}$ is a submatrix of B. We will now show that $\widetilde{x}$ is a nondegenerate basic feasible solution, so that in fact $B = A_+$. Since $A\widetilde{x} = b + \sum_{j=1}^n A_j \epsilon^j$, and $\widetilde{x}_j = 0$, $j \in N$, it holds that

$$0 \le e_i^T \widetilde{x}_B = e_i^T B^{-1} b + \epsilon^{B(i)} + \sum_{j\in N} e_i^T B^{-1} A_j \epsilon^j, \quad i = 1, \dots, m, \tag{4.6}$$

where $B(i)$ denotes the index of the ith basic variable. From our choice of ϵ, Lemma A.1 in conjunction with (4.6) shows that $e_i^T \widetilde{x}_B > 0$, $i = 1, \dots, m$, so that $\widetilde{x}$ is a nondegenerate optimal basic feasible solution. Lemma 4.1 shows that there are y and s such that $y = B^{-T} c_B$, $s_B = 0$, $s_N = c_N - N^T y$ and $s_N \ge 0$, i.e., (4.4c) and (4.4d) hold. In addition, Lemma A.1 shows that if $e_i^T B^{-1} b < 0$ for some i, then (4.6) cannot hold. Hence, $B^{-1} b \ge 0$. We may thus define x by $x_B = B^{-1} b$, $x_N = 0$, so that (4.4a) and (4.4b) hold. Since x, y and s satisfy (4.4), Proposition 2.1 shows that x is optimal to (LP). ∎

Proposition 4.1 ensures the existence of an optimal basic feasible solution for which we may associate a basis matrix such that (4.4) holds. As shown in the following corollary, we may find at least one such basis matrix for each optimal basic feasible solution.

Corollary 4.1. *Assume that (LP) is feasible, and that there is no $p \in I\!R^n$ such that $p \ge 0$, $Ap = 0$, $c^T p < 0$. Then, for an optimal basic feasible solution $x \in I\!R^n$,* *there is an associated basis matrix B and vectors $y \in I\!R^m$ and $s \in I\!R^n$ such that*

$$\begin{aligned} &Bx_B = b, \quad x_N = 0, && (4.7a)\\ &x_B \geq 0, && (4.7b)\\ &B^T y = c_B, \quad s_B = 0, \quad N^T y + s_N = c_N, && (4.7c)\\ &s_N \geq 0. && (4.7d) \end{aligned}$$

Proof. Let x be an optimal basic feasible solution. It follows from Proposition 4.1 that there exists an optimal basic feasible solution $\widetilde{x}$ and associated basis matrix $\tilde{B}$ in addition to vectors y and s such that

$$\begin{aligned} &\textstyle\sum_{j\in\tilde{B}} A_j \widetilde{x}_j = b, \quad \widetilde{x}_j \geq 0, \quad j \in \tilde{B}, && (4.8a)\\ &A_j^T y = c_j, \quad j \in \tilde{B}, && (4.8b)\\ &A_j^T y + s_j = c_j, \quad s_j \geq 0, \quad j \in \widetilde{N}, && (4.8c) \end{aligned}$$

Since $s^T\widetilde{x} = 0$, and x is optimal, Proposition 2.1 implies that $s^T x = 0$ must hold. Consequently, it follows that $s_j = 0$, $j \in A_+$, and we conclude that

$$A_j^T y = c_j, \quad j \in \tilde{B} \cup A_+. \tag{4.9}$$

As $\tilde{B}$ is nonsingular and A_+ has full column rank, we may create a nonsingular matrix B by adding columns of $\tilde{B}$ to A_+. For this B, (4.7a) and (4.7b) will hold, since A_+ is a submatrix of B and x is feasible to (LP). By (4.8c) and (4.9), it follows that (4.7c) and (4.7d) hold. ∎

5. Necessary and sufficient optimality conditions

Given the existence of y and s, which was shown for an optimal basic feasible solution, we may derive necessary optimality conditions for an arbitrary point, which is not necessarily a basic feasible solution.

Corollary 5.1. *Assume that (LP) is feasible, and that there is no $p \in I\!R^n$ such that $p \geq 0$, $Ap = 0$, $c^T p < 0$. Then, (LP) has at least one optimal solution. A vector $x \in I\!R^n$ is optimal to (LP) if and only if there are $y \in I\!R^m$, $s \in I\!R^n$ such that*

$$\begin{aligned} Ax &= b, && (5.1a)\\ x &\geq 0, && (5.1b)\\ A^T y + s &= c, && (5.1c)\\ s &\geq 0, && (5.1d)\\ s^T x &= 0. && (5.1e) \end{aligned}$$

The optimal value is given by $b^T y$.

Proof. As for sufficiency, Proposition 2.1 shows that if x, y and s satisfy (5.1), then x is optimal to (LP).

As for necessity, Proposition 4.1 ensures that if there is no $p \in I\!R^n$ such that $p \geq 0$, $Ap = 0$, $c^T p < 0$, there is at least one optimal x for which there exist y and s such that (5.1) holds. Let $\widetilde{x}$ be any optimal solution to (LP). Then $\widetilde{x}$ must satisfy (5.1a) and (5.1b) to be feasible to (LP). In addition, Proposition 2.1 shows that the optimal value of (LP) is $b^T y$ and in addition that $\widetilde{x}$ is optimal only if $s^T \widetilde{x} = 0$. Hence, (5.1e) holds too. ∎

6. A proof of Farkas' lemma

Although the purpose of this note is not to present a proof of Farkas' lemma, we observe that Farkas' lemma may be derived directly from the optimality conditions.

Proposition 6.1. (Farkas' lemma) *Let M be an $m \times n$ matrix, and let g be an m-dimensional vector. Then exactly one of the following systems has a solution:*

$$1. \quad Mu = g, \; u \geq 0, \qquad\qquad 2. \quad M^T y \leq 0, \; g^T y > 0.$$

Proof. We first show that 1 and 2 cannot have a solution simultaneously. This is done by contradiction. Assume that u satisfies 1 and y satisfies 2. Then, premultiplication of both sides of $Mu = g$ by y^T gives $y^T M u = y^T g$. By 2, $y^T g > 0$. On the other hand, $u \geq 0$ and $M^T y \leq 0$ gives $y^T M u \leq 0$. Hence, we have a contradiction, showing that 1 and 2 cannot have a solution simultaneously.

It now suffices to show that one of 1 and 2 always has a solution. For this purpose, consider the linear program

$$\begin{array}{ll} \underset{u \in I\!R^n, w \in I\!R^m}{\text{minimize}} & e^T w \\ \text{subject to} & Mu + Dw = g, \\ & u \geq 0, \quad w \geq 0, \end{array} \tag{6.1}$$

where e is a vector of ones and D is a diagonal matrix with diagonal elements $d_{ii} = 1$ if $g_i \geq 0$ and $d_{ii} = -1$ if $g_i < 0$, for $i = 1, \ldots, m$. Note that the linear program (6.1) is feasible, since $u = 0$, $w = |g|$ gives a feasible solution. The constraint matrix has full row rank, since D is nonsingular. In addition, the optimal value is bounded below by zero, since the nonnegativity of w implies that the objective function is the 1-norm of w. Hence, Lemma 3.1 implies that there exists at least one optimal solution. There are two cases, either the optimal value is zero, or it is strictly positive.

If the optimal value of (6.1) is zero, let u and w denote an optimal solution. It must hold that $e^T w = 0$, $w \geq 0$, i.e., $w = 0$. Hence, u satisfies $Mu = g$, $u \geq 0$, which means that u fulfills 1.

If the optimal value of (6.1) is strictly positive, Corollary 5.1 implies that there are y, s_1 and s_2 such that

$$M^T y + s_1 = 0, \quad s_1 \geq 0, \tag{6.2a}$$

$$Dy + s_2 = e, \quad s_2 \geq 0. \tag{6.2b}$$

with $g^T y > 0$. In particular, (6.2a) gives $M^T y \le 0$. Hence, y fulfills 2. ∎

7. Conclusion

We have given an elementary proof of optimality conditions for linear programming. The advantage compared to using the simplex method for giving these conditions is that we have not needed to introduce the simplex method. The advantage compared to using Farkas' lemma is that we have only introduced a perturbation of the constraints, which can be interpreted directly in the original problem. An additional advantage to using Farkas' lemma is that our approach naturally gives the existence of an associated basis matrix for each optimal basic feasible solution which may be used to give the optimality conditions. To keep the exposition simple, we have not introduced the concept of duality. However, for linear programming the optimality conditions are essentially equivalent to strong duality. Hence, the results given also prove strong duality.

A. A perturbation lemma

For completeness, we give the following perturbation result which shows that the sign of a polynomial $\sum_{j=0}^n a_j \epsilon^j$, for ϵ positive and sufficiently small, is determined by the sign of the first nonzero coefficient a_j, with the ordering $j = 0, 1, \ldots, n$. See, e.g., Dantzig [3, Lemma 1, Chapter 10] for a similar result.

Lemma A.1. *Let $a \in I\!R^n$, $a \neq 0$, and let γ and Γ be positive numbers such that $0 < \gamma \le \min_{j:a_j \neq 0} |a_j|$ and $\Gamma \ge \max_{j:a_j \neq 0} |a_j|$. Since $a \neq 0$, there is an index k such that $a_j = 0$, $j = 0, \ldots, k-1$, $a_k \neq 0$. If $0 < \epsilon < \gamma/(\gamma + \Gamma)$ then*

$$\sum_{j=0}^n a_j \epsilon^j > 0 \text{ if } a_k > 0,$$

$$\sum_{j=0}^n a_j \epsilon^j < 0 \text{ if } a_k < 0.$$

Proof. Assume that $a_k > 0$, $a_j = 0$, $j = 0, \ldots, k-1$, $0 < \gamma \le \min_{j:a_j \neq 0} |a_j|$ and $\Gamma \ge \max_{j:a_j \neq 0} |a_j|$, with $\gamma < \Gamma$. If $0 < \epsilon < \gamma/(\gamma + \Gamma)$, then

$$\begin{aligned}
\sum_{j=0}^n a_j \epsilon^j &= a_k \epsilon^k + \sum_{j=k+1}^n a_j \epsilon^j = \epsilon^k \left(a_k + \sum_{j=k+1}^n a_j \epsilon^{j-k} \right) \\
&\ge \epsilon^k \left(\gamma - \Gamma \sum_{j=1}^{n-k} \epsilon^j \right) \ge \epsilon^k \left(\gamma - \Gamma \sum_{j=1}^{\infty} \epsilon^j \right) = \epsilon^k \left(\gamma - \Gamma \frac{\epsilon}{1-\epsilon} \right) \\
&= \epsilon^k \left(\frac{\gamma - (\Gamma + \gamma)\epsilon}{1 - \epsilon} \right) > 0,
\end{aligned}$$

where the last inequality follows from the choice of ϵ. The proof for the case when $a_k < 0$ is analogous. ∎

Acknowledgement

This note has its origin in discussions with Krister Svanberg on elementary proofs of Farkas' lemma.

References

[1] A. Charnes. Optimality and degeneracy in linear programming. *Econometrica*, 20(2):160–170, apr 1952.

[2] V. Chvátal. *Linear Programming.* W. H. Freeman and Company, New York, 1983.

[3] G. B. Dantzig. *Linear Programming and Extensions.* Princeton University Press, Princeton, New Jersey, 1963.

[4] A. Dax. An elementary proof of Farkas' lemma. *SIAM Rev.*, 39(3):503–507, 1997.

[5] P. E. Gill, W. Murray, and M. H. Wright. *Numerical Linear Algebra and Optimization, volume 1.* Addison-Wesley Publishing Company, Redwood City, 1991. ISBN 0-201-12649-4.

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

[7] L. G. Khachian. A polynomial algorithm in linear programming. *Dokl. Akad. Nauk SSSR*, 244(5):1093–1096, 1979. Translated in *Soviet Math. Doklady*, 20:191–194, 1979.

[8] D. G. Luenberger. *Linear and Nonlinear Programming.* Addison-Wesley Publishing Company, Reading, second edition, 1984. ISBN 0-201-15794-2.

[9] S. G. Nash and A. Sofer. *Linear and Nonlinear Programming.* McGraw-Hill, New York, 1996. ISBN 0-07-046065-5.

[10] A. Schrijver. *Theory of Linear and Integer Programming.* John Wiley and Sons, Inc., New York, Chichester, Brisbane, Toronto and Singapore, 1986.

[11] K. Svanberg. Farkas' lemma derived by elementary linear algebra. Manuscript, Royal Institute of Technology, 2008.