
An Extension of the Fundamental Theorem of Linear Programming

A. Brown^{a*}, A. Gedlaman^{b*}, A. Holder^{b*†}, S. Martinez^{c*}

^a Department of Mathematics, Saint Mary's University, San Antonio, TX.

^b Department of Mathematics, Arizona State University, Tempe, AZ.

^c Hearin Center for Enterprise Science, School of Business Administration, The University of Mississippi. aholder@trinity.edu

^d Department of Mathematics, University of Texas at San Antonio, TX.

Latest Revision: September 2001

Abstract — In 1947 George Dantzig developed the Simplex Algorithm for linear programming, and in doing so became known as The Father of Linear Programming. The invention of the Simplex Algorithm has been called “one of the most important discoveries of the 20th century,” and linear programming techniques have proven useful in numerous fields of study. As such, topics in linear optimization are taught in a variety of disciplines. The finite convergence of the simplex algorithm hinges on a result stating that every linear program with an optimal solution has a basic optimal solution; a result known as the Fundamental Theorem of Linear Programming. We develop an analog to the fundamental theorem, and the perspective from which we view the problem allows a much greater class of functions. Indeed, not only do we relinquish the assumption of linearity, but we also do not assume the functions under consideration are continuous. Our new result implies the Fundamental Theorem of Linear Programming.

Keywords: Linear Programming.

* Research supported by NSF grant 0097366. Research conducted at Trinity University.

† Research supported by ONR grant N00014-01-1-0917. Research conducted at The University of Mississippi.

1 Introduction

The field of linear programming began when George Dantzig invented the simplex algorithm in 1947, and this area of mathematics has influenced numerous disciplines (Business, Health Care, Computer Science, Economics, and Military Science to name just a few). Today, linear programming topics are taught within mathematics, business, economics, and computer science courses, and the topics most commonly covered are problem formulation, geometric interpretation, and a tableau presentation of the simplex method. The motivation and finite convergence of the simplex method rely on the Fundamental Theorem of Linear Programming, a result that we briefly review and extend after discussing some notation.

We use \mathbb{R}_+^n and \mathbb{R}_{++}^n to indicate the collection of vectors in \mathbb{R}^n with non-negative and positive components, respectively. Furthermore, a subscript of L indicates that we are imposing the lexicographic ordering —i.e. \mathbb{R}^n and \mathbb{R}_L^n are the same set, but the vectors in \mathbb{R}_L^n are lexicographically ordered. Let x^1 and x^2 be in \mathbb{R}_L^n . If $x^1 \neq x^2$, let k be such that it is the first index such that $x_k^1 \neq x_k^2$. We have that $x^1 <_L x^2$ if $x_k^1 < x_k^2$, and that $x^1 \leq_L x^2$ if $x^1 <_L x^2$ or $x^1 = x^2$. An epsilon neighborhood of x is denoted by $\mathcal{N}_\varepsilon(x)$. Overline notation is used to indicate the closure of a set, so $\overline{\mathcal{N}_\varepsilon(x)} = \{y : \|x - y\| \leq \varepsilon\}$. A set subscript on a vector indicates the subvector whose elements correspond to the elements in the set. Similarly, a set subscript on a matrix indicates the submatrix whose columns correspond to the indices in the set. Furthermore, for $x \in \mathbb{R}_+^n$ we let $B(x) = \{i : x_i > 0\}$ and $N(x) = \{i : x_i = 0\}$, where we do not include the argument when it is clear which x is under consideration. As an example, we have

$$Ax = \begin{bmatrix} 1 & 2 & -1 \\ 0 & 1 & -1 \end{bmatrix} \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix} = \begin{bmatrix} 1 & -1 \\ 0 & -1 \end{bmatrix} \begin{pmatrix} 1 \\ 2 \end{pmatrix} + \begin{bmatrix} 2 \\ 1 \end{bmatrix} \begin{pmatrix} 0 \end{pmatrix} = A_B x_B + A_N x_N = A_B x_B.$$

The Null Space of the matrix A is denoted by $\text{Nul}(A)$. The difference between a term in bold face and a term in italics is that the former is defined within the context of this work and the later is considered common knowledge (and is found in the *Mathematical Programming Glossary* [1]).

Consider the linear program, $(LP) \min\{c^T x : Ax = b, x \geq 0\}$, where $c \in \mathbb{R}^n$, $b \in \mathbb{R}^m$, and $A \in \mathbb{R}^{m \times n}$. We assume without loss of generality that $m \leq n$, and that A has full row rank. The **feasible region**, which we denote by $\mathcal{P} = \{x : Ax = b, x \geq 0\}$, is a polyhedron formed by taking the *affine space* $\{x : Ax = b\}$ and intersecting it with the positive orthant \mathbb{R}_+^n . As an example consider the linear program $\min\{c^T x : [1, 2, 3]x = 3, x \geq 0\}$, whose feasible region is depicted in Figure 1. The feasible region for this example has 3 *extreme points*: $(3, 0, 0)^T$, $(0, 3/2, 0)^T$, and $(0, 0, 1)^T$.

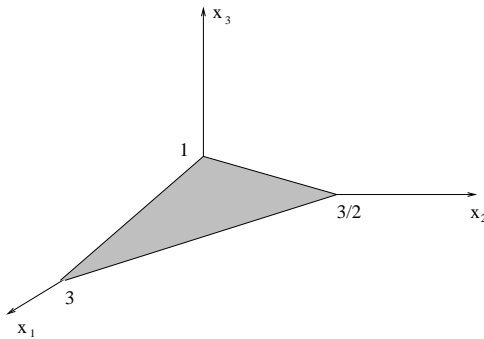


Figure 1: The feasible region for $\min\{c^T x : [1, 2, 3]x = 3, x \geq 0\}$.

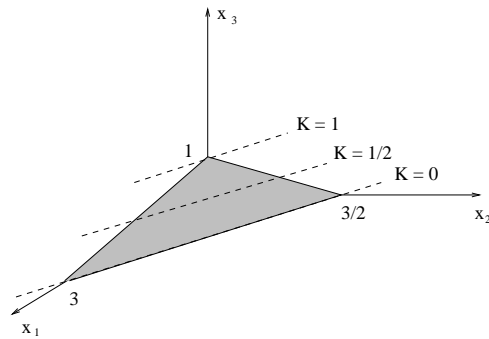


Figure 2: Level sets for $\min\{(0, 0, 1)^T x : [1, 2, 3]x = 3, x \geq 0\}$.

The geometric insight into the Fundamental Theorem of Linear Programming comes from considering **level sets** of the objective function. For $K \in \mathbb{R}$, the K^{th} level set is $\mathcal{L}_K = \{x : c^T x = K\}$, and a solution to (LP) is the minimum K such that $\mathcal{L}_K \cap \mathcal{P} \neq \emptyset$. An example is depicted in Figure 2,

where $c = (0, 0, 1)^T$ and any feasible element with $x_3 = 0$ is optimal. An important observation is that no matter which c is used in this example, there always exists an extreme point of \mathcal{P} that is optimal. For the example in Figure 2 we have that both $(3, 0, 0)^T$ and $(0, 3/2, 0)^T$ are optimal.

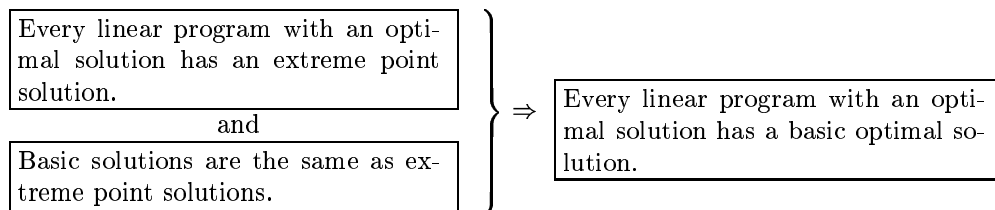
From this geometric insight we see that to solve (LP) we do not need to investigate every feasible element, but rather only the extreme points of the feasible region. This means that we need a convenient manner to identify the extreme points of a polyhedron. The following definition defines a **basic solution**, and Lemma 1 shows that extreme points and basic solutions are one and the same.

Definition 1 (Basic Solution) *The vector $x \in \mathcal{P}$ is a **basic solution** if the columns of A_B are linearly independent.*

Each basic solution corresponds to at least one basis for \mathbb{R}^n . If x is a basic solution and $A_{B(x)}$ is square, and hence invertible, the columns of $A_{B(x)}$ form a unique basis. In this case the basic solution is **non-degenerate**. Otherwise, columns of $A_{N(x)}$ are needed to extend the collection of columns of $A_{B(x)}$ to a basis, and the basis is not necessarily unique. Such basic solutions are called **degenerate**.

Lemma 1 (See [2] for proof) *A feasible element is basic if, and only if, it is an extreme point of the feasible region.*

Lemma 1 provides an algebraic description of an extreme point, and the simplex algorithm uses this connection to search the extreme points of the feasible region. While a complete discussion of the simplex method is beyond the scope of this work, and is not required to understand our new result, the following logical sequence of thought is important.



This conclusion is known as the Fundamental Theorem of Linear Programming.

Theorem 1 (See [2] for proof) *If a linear program has an optimal solution, then there exists a basic optimal solution.*

The main objective of this work is to provide a considerably weaker set of conditions that provide an analog to the fundamental theorem. In fact, Theorem 1 follows directly from our new result. Section 2 motivates and develops the weaker conditions needed to prove our extension. An example is provided to show the generality with which our result holds. The section concludes by showing how the Fundamental Theorem of Linear Programming follows from our result.

2 An Analog to The Fundamental Theorem of Linear Programming

In this section we consider the multiple objective program (MOP) $\min\{F(x) : Ax = b\}$, where $F : \mathbb{R}^n \rightarrow \mathbb{R}_L^p$ and $A \in \mathbb{R}^{m \times n}$ has full row rank. We denote the feasible region by \mathcal{P} and the optimal set by \mathcal{O} . We further make the tacit assumption that $\text{Nul}(A) \neq \{0\}$, for if it were, the feasible region would contain only a single element. The goal of this section is to develop properties for F that permit an extension of Theorem 1. The properties fall into two categories. First, we need a way to define an object similar to a basic solution. The difficulty here is that the minimization for (MOP) is over an affine space —i.e. there are no inequality constraints. This means that the index sets $B(x)$ and $N(x)$ do not make sense. Notice that these index sets rely only on the feasible element x and the coefficient matrix A , and are defined independent of the objective function. Our

approach is different in that the index sets we develop depend on the objective function. These new index sets allow us to define a **corner**, which is an extension of the idea of a basic solution (for the case of piecewise linear functions see [4]). Second, we require that the objective function in (MOP) has a monotonicity property over line segments passing through the optimal set. This condition is precisely defined in Definition 3.

We assume that F has the following form,

$$\begin{aligned} F(x) &= (F_1(x), F_2(x), \dots, F_p(x))^T, \text{ where} \\ F_i(x) &= \sum_{j=1}^n f_{(i,j)}(x_j), \text{ for } i \in \{1, 2, \dots, p\} \text{ and } x \in \mathbb{R}^n. \end{aligned}$$

Observe that $F_i(x)$ is **separable**, meaning that $f_{(i,j)}$ has only x_j as its argument.

We describe the behavior of a function “near” a point by saying that a function exhibits a behavior **locally at a point** if there exists a neighborhood of that point over which that behavior holds. For example, a real-valued function f is locally monotonic at x if there exists a neighborhood of x over which f is monotonic. Similarly, f is locally constant at x if there exists a neighborhood of x over which f is constant. We also say that a real-valued function is **strongly monotonic** over an interval if it is either strictly increasing, strictly decreasing, or constant.

Unless $n = p = 1$, the concept of monotonicity does not make sense for F . However, we use the property that F is separable to capture, in some sense, where F changes monotonicity over each axis. We examine those points at which F “changes monotonicity” by defining

$$H_j = \{x_j : f_{(i,j)} \text{ is not locally strongly monotonic at } x_j \text{ for some } i \in \{1, 2, \dots, p\}\}.$$

An important observation is that each H_j is closed. This follows because for an element to be in the complement of H_j means that $f_{(i,j)}$ is locally strongly monotonic for all $i \in \{1, 2, \dots, p\}$. Hence, there is an open neighborhood about the element for which each $f_{(i,j)}$ remains locally strongly monotonic.

As an example, suppose that \mathbb{Q} is the set of rationals and that $F : \mathbb{R}^2 \rightarrow \mathbb{R}_L^3$ is defined by

$$F(x) = (F_1(x), F_2(x), F_3(x))^T = \left(\sum_{j=1}^2 f_{(1,j)}(x_j), \sum_{j=1}^2 f_{(2,j)}(x_j), \sum_{j=1}^2 f_{(3,j)}(x_j) \right)^T,$$

where

$$\begin{aligned} f_{(1,1)}(x_1) &= x_1 + \sin x_1, \\ f_{(1,2)}(x_2) &= (x_2)^3, \\ f_{(2,1)}(x_1) &= \begin{cases} 5, & x_1 < 1 \\ 0, & x_1 \geq 1, x_1 \in \mathbb{Q} \\ 1, & \text{otherwise,} \end{cases} \\ f_{(2,2)}(x_2) &= -x_2, \\ f_{(3,1)}(x_1) &= e^{x_1}, \text{ and} \\ f_{(3,2)}(x_2) &= (x_2 - 1)^2. \end{aligned}$$

Notice that $f_{(1,1)}$ and $f_{(3,1)}$ are strictly increasing. Thus, they do not contribute any points to H_1 . The function $f_{(2,1)}$ is constant over $(-\infty, 1)$, but for $x_1 \geq 1$ there is no point at which $f_{(2,1)}$ is locally strongly monotonic. Thus $f_{(2,1)}$ contributes the entire interval $[1, \infty)$ to H_1 . We have that $f_{(1,2)}$ and $f_{(2,2)}$ contribute no points to H_2 because they are strictly monotonic. However, $f_{(3,2)}$ is not locally strongly monotonic at $x_2 = 1$, and hence, $H_2 = \{1\}$.

A component x_j is **cornered** if $x_j \in H_j$. The following index sets indicate which components are cornered and which are not,

$$\begin{aligned} \nu(x) &= \{j : x_j \text{ is cornered}\}, \text{ and} \\ \beta(x) &= \{j : x_j \text{ is not cornered}\}. \end{aligned}$$

Notice that $\nu(x)$ and $\beta(x)$ are similar to $N(x)$ and $B(x)$, the difference being that N and B indicate which components of x are 0 and are not 0, while ν and β indicate which components are cornered and are not cornered. The index sets $\nu(x)$ and $\beta(x)$ allow us to define a corner in the following manner.

Definition 2 (Corner Solution) *The vector $x \in \mathcal{P}$ is a **corner solution** if the columns of $A_{\beta(x)}$ are linearly independent.*

The set of corners is denoted by \mathcal{C} . As before, a corner is non-degenerate if $A_{\beta(x)}$ is invertible, and is otherwise degenerate. Notice that the definition of a corner is similar to the definition of a basic solution, the difference being in how the index sets are defined. Returning to our previous example of F , suppose that the constraint $Ax = 0$ is included. If $A = [1, 0]$ the feasible region is the x_2 axis. Since $0 \notin H_1$, x_1 cannot be cornered. The component x_2 is cornered only if $x_2 = 1$, since that is the only point in H_2 . Since $\beta((0, 1)^T) = \{1\}$, and $A_{\beta((0, 1)^T)} = [1]$, we have that $x = (0, 1)^T$ is a corner (and is in fact the only one). If $A = [1, -1]$, the feasible region is the line $x_1 = x_2$, and we have corners at every point $(x', x')^T$, so long as $x' \geq 1$. If $x' = 1$, $\beta((x', x')^T) = \emptyset$, and $A_{\beta(x', x')^T}$ is vacuous (and hence its columns are linearly independent). For this case we have that (x', x') is a degenerate corner. If $x' > 1$, $\beta((x', x')^T) = \{2\}$ and $A_{\beta((x', x')^T)} = [-1]$, from which we have that $(x', x')^T$ is a non-degenerate corner.

Since our goal is to prove a result about the existence of a corner optimal solution, we need to require that a corner exists. Hence, we make the following assumption.

Assumption 1 *We assume that F has the property that $H_j \neq \emptyset$, for all $j \in \{1, 2, \dots, n\}$.*

The full row rank of A implies that, with a suitable reordering of the columns, A may be partitioned as $A = [A' | A'']$, where A' is invertible. Partitioning any feasible x appropriately, we have $Ax = A'x' + A''x''$. From Assumption 1 we have that x'' may be chosen so that each of its components is cornered. Setting $x' = (A')^{-1}(b - A''x'')$ provides us with a corner solution. So, Assumption 1, and the fact that A has full row rank, guarantee the existence of a corner.

Other than making sure a corner exists, we need for F to have a monotonicity property along arbitrary line segments. Open and closed line segments are, respectively, denoted by

$$\begin{aligned} l(x^1, x^2) &= \{x : x = (1 - \theta)x^1 + \theta x^2, \theta \in (0, 1)\}, \text{ and} \\ \bar{l}(x^1, x^2) &= \{x : x = (1 - \theta)x^1 + \theta x^2, \theta \in [0, 1]\}. \end{aligned}$$

For an understood x^1 and x^2 , we make the notational convention that $x(\theta) = (1 - \theta)x^1 + \theta x^2$. We state that f is **monotonic over** a line segment if $\theta_1 \leq \theta_2$ implies either $f(x(\theta_1)) \geq f(x(\theta_2))$ or $f(x(\theta_1)) \leq f(x(\theta_2))$, for all θ_1 and θ_2 allowed in the definition of the line segment. Similarly, f is **strictly monotonic over** a line segment if $\theta_1 < \theta_2$ implies either $f(x(\theta_1)) < f(x(\theta_2))$ or $f(x(\theta_1)) > f(x(\theta_2))$. We use this concept of monotonicity over a line segment to define what it means for a function to be **linearly monotonic**.

Definition 3 (Linear Monotonicity) *A real-valued function f is (strongly) linearly monotonic if it is (strongly) monotonic on $\bar{l}(x^1, x^2)$ for all $x^1, x^2 \in R^n$. f is (strongly) linearly monotonic over S if it is (strongly) monotonic over $l(x^1, x^2)$ for all $x^1, x^2 \in S$ and $l(x^1, x^2) \cap S = l(x^1, x^2)$.*

Definition 3 embodies a property that we require of F . We are interested in line segments that intersect the optimal set with the property that all points on the line segment have the same cornered components. To formally define the subset of interest, we define

$$Q(x) = \{x + \alpha q : \alpha \in \mathbb{R}, q \in \text{Nul}(A), \nu(x + \alpha q) = \nu(x), x_{\nu(x)} = (x + \alpha q)_{\nu(x)}\}.$$

The set $Q(x)$ is the collection of feasible elements that 1) are on a line segment with x , 2) the index set of cornered components is the same for each point on the line segment, and 3) the cornered components have the same value along the line segment. The third condition indicates that $q_{\nu(x)}$ must be zero for $\alpha \neq 0$. For $x \in \mathcal{O}$, assumption 2 uses the line segments in $\bar{Q}(x)$ to describe the monotonicity property that we require.

Assumption 2 For each $x \in \mathcal{O} \setminus \mathcal{C}$, we assume that $F(x) = (F_1(x), F_2(x), \dots, F_p(x))^T$ has the property that each $F_i(x)$, $i = 1, 2, \dots, p$, is strongly linearly monotonic over $\overline{Q}(x)$.

Assumption 2 asserts that the component functions of F are strongly monotonic over line segments that contain a non-corner optimal element, provided that the cornered components adhere to the restrictions in $Q(x)$. We exclude the corners from the optimal set because we gain some generality.

The following lemma shows we can increase the number of cornered components of a non-corner optimal solution.

Lemma 2 If x^0 is a non-corner optimal solution, there exists x^1 such that

- (i) $x^1_{\nu(x^0)} = x^0_{\nu(x^0)}$,
- (ii) $\nu(x^0) \subset \nu(x^1)$, and
- (iii) x^1 is optimal.

Proof: Note that since x^0 is not a corner, the columns of $A_{\beta(x^0)}$ are linearly dependent. It follows that $\text{Nul}(A_{\beta(x^0)}) \neq \{0\}$. Choose a direction vector

$$q = \left[q_{\beta(x^0)}^T \mid q_{\nu(x^0)}^T \right]^T, \text{ where } q_{\beta(x^0)} \in \text{Nul}(A_{\beta(x^0)}) \setminus \{0\} \text{ and } q_{\nu(x^0)} = 0.$$

For $\alpha \in \mathbb{R}$, let $x(\alpha)$ be the partitioned column vector

$$x(\alpha) = x^0 + \alpha q = \begin{bmatrix} x_{\beta(x^0)}^0 + \alpha q_{\beta(x^0)} \\ x_{\nu(x^0)}^0 \end{bmatrix}.$$

We have that $x(\alpha)$ is feasible because

$$\begin{aligned} Ax(\alpha) &= A(x^0 + \alpha q) \\ &= A_{\beta(x^0)}[x_{\beta(x^0)}^0 + \alpha q_{\beta(x^0)}] + A_{\nu(x^0)}x_{\nu(x^0)}^0 \\ &= \left(A_{\beta(x^0)}x_{\beta(x^0)}^0 + A_{\nu(x^0)}x_{\nu(x^0)}^0 \right) + \alpha A_{\beta(x^0)}q_{\beta(x^0)} \\ &= Ax^0 + \alpha A_{\beta(x^0)}q_{\beta(x^0)} \\ &= b. \end{aligned}$$

We now find the smallest positive α that will corner a component of $x_{\beta(x^0)}$. However, we first need to ensure that there exists $j \in \beta(x^0)$ such that q directs x_j^0 towards some $h \in H_j$. Suppose that for all $j \in \beta(x^0)$, either $x_j^0 > \sup\{H_j\}$ and $q_j \geq 0$, or $x_j^0 < \inf\{H_j\}$ and $q_j \leq 0$. Then, for all $j \in \beta(x^0)$ and $\alpha \geq 0$, $x_j^0 + \alpha q_j \notin H_j$. Thus it is impossible to corner any component of $x_{\beta(x^0)}$. However, since $q_{\beta(x^0)} \in \text{Nul}(A_{\beta(x^0)})$, we also know that $-q_{\beta(x^0)} \in \text{Nul}(A_{\beta(x^0)})$. So, if q does not direct x_j^0 towards $h \in H_j$, then $-q$ does. Hence, we may always choose $q \in \text{Nul}(A)$, with $q_{\beta(x^0)} \in \text{Nul}(A_{\beta(x^0)})$ and $q_{\nu(x^0)} = 0$, such that q directs x_j^0 towards some $h \in H_j$, for some $j \in \beta(x^0)$. We assume throughout that q has this property.

For notational brevity, we conveniently define $\min \emptyset$ to be ∞ . This notational convenience, together with the fact that each H_j is closed, allows us to use the min operator to define

$$\begin{aligned} \hat{\alpha}_+ &= \min \left\{ \frac{h - x_j^0}{q_j} : j \in \beta(x^0), h \in H_j, q_j > 0, h > x_j^0 \right\}, \text{ and} \\ \hat{\alpha}_- &= \min \left\{ \frac{h - x_j^0}{q_j} : j \in \beta(x^0), h \in H_j, q_j < 0, h < x_j^0 \right\}. \end{aligned}$$

Let $\hat{\alpha} = \min \{\hat{\alpha}_+, \hat{\alpha}_-\}$ (notice that $\hat{\alpha}$ is positive), and set $x^1 = x(\hat{\alpha}) = x^0 + \hat{\alpha}q$. Observe that for all $j \in \nu(x^0)$, we have that $x_j^1 = x_j^0 + \hat{\alpha}q_j = x_j^0$. We conclude that $x^1_{\nu(x^0)} = x^0_{\nu(x^0)}$, which proves statement (i).

By the definition of x^1 , we know there exists $j \in \beta(x^0)$ such that $x_j^1 \in H_j$ and $x_j^0 \notin H_j$. So, j is in $\nu(x^1)$ but not in $\nu(x^0)$. Since $x_{\nu(x^0)}^1 = x_{\nu(x^0)}^0$, we have that $\nu(x^0) \subset \nu(x^1)$, and statement (ii) is proven.

From the definition of $\hat{\alpha}$ we have that $x(\alpha) \in Q(x^0)$, for $\alpha \in [0, \hat{\alpha}]$. Since each F_i is strongly linearly monotonic over $\overline{Q}(x^0)$, we have that each $F_i(x(\alpha))$ is strongly monotonic over $[0, \hat{\alpha}]$. From the fact that each H_j is closed we have that for each $j \in \beta(x^0)$, there exists a neighborhood, $\mathcal{N}_{\varepsilon_j}(x_j^0)$, such that $\mathcal{N}_{\varepsilon_j}(x_j^0) \cap H_j = \emptyset$. Let $\bar{\alpha} = \min \left\{ \frac{\varepsilon_j}{|q_j|} : j \in \beta(x^0), q_j \neq 0 \right\}$. Then, for all $\alpha \in (-\bar{\alpha}, \bar{\alpha})$, and all $j \in \beta(x^0)$, $x_j(\alpha) \notin H_j$. This, together with the fact that $x_{\nu(x^0)}(\alpha)$ is constant for all α , implies that $x(\alpha) \in Q$, for $\alpha \in (-\bar{\alpha}, \bar{\alpha})$. We now have that $l(x(-\bar{\alpha}), x(\hat{\alpha}))$ is contained in $Q(x^0)$. Since each F_i is strongly linearly monotonic over $\overline{Q}(x^0)$, we have that each $F_i(x(\alpha))$ is strongly monotonic over $[-\bar{\alpha}, \bar{\alpha}]$.

Since x^0 is optimal, we conclude that $F(x^0) \leq_L F(x^1)$. Suppose for the sake of attaining a contradiction that $F(x^0) <_L F(x^1)$. Let $k \in \{1, 2, \dots, p\}$ be the smallest index such that $F_k(x^0) < F_k(x^1)$. First, since $F_i(x^0) = F_i(x^1)$, for $i < k$, and each $F_i(x(\alpha))$ is strongly monotonic over $[-\bar{\alpha}, \bar{\alpha}]$, we conclude that $F_i(x(\alpha))$ is constant over $[-\bar{\alpha}, \bar{\alpha}]$. Second, since $F_k(x^0) < F_k(x^1)$ and $F_k(x(\alpha))$ is strongly monotonic over $[-\bar{\alpha}, \hat{\alpha}]$, we have that $F_k(x^0) > F_k(x^0 - \bar{\alpha}q)$. However, this contradicts the fact that x^0 is optimal. So, $F(x^0) = F(x^1)$, and statement (iii) is proven. ■

Theorem 2 *If $\min\{F(x) : Ax = b\}$ has a solution, then it has a corner optimal solution.*

Proof: Let x^0 be an optimal solution. If x^0 is a corner, then the proof is complete. If x^0 is not a corner, Lemma 2 implies that there exists an optimal solution x^1 such that $x_{\nu(x^0)}^1 = x_{\nu(x^0)}^0$ and $\nu(x^0) \subset \nu(x^1)$. This means that $A_{\beta(x^1)}$ is formed by removing columns from $A_{\beta(x^0)}$. There are two possible cases. First, if the columns of $A_{\beta(x^1)}$ are linearly independent, x^1 is a corner optimal solution, and we are done. Second, if the columns of $A_{\beta(x^1)}$ are linearly dependent, we apply Lemma 2 again to get an optimal x^2 such that $A_{\beta(x^2)}$ is formed by removing columns from $A_{\beta(x^1)}$. In this manner we form a sequence of optimal solutions, say x^1, x^2, \dots, x^k , such that the columns of $A_{\beta(x^k)}$ are linearly independent (this matrix is empty if $\beta(x^k) = \emptyset$). Hence, x^k is an optimal corner. ■

The following example illustrates the complexity of the functions allowed by Theorem 2

Example 1 *Let $A = [0, 0, 1]$ and $b = [0]$ so that the feasible region is $\mathbb{R}^2 \times \{0\}$. Let $F_1(x) = \sum_{i=1}^3 f_{(1,j)}(x_j)$, where each term is defined by*

$$f_{(1,j)}(x_j) = \begin{cases} 1, & x_j < 0 \\ 0, & x_j \geq 0. \end{cases}$$

Notice that the definition of $f_{(1,j)}$ implies that 0 is in H_j , for $j = 1, 2, 3$. So, Assumption 1 is satisfied.

The second objective function is purposely more complicated so as to demonstrate the robustness of Theorem 2. Since every feasible element has $x_3 = 0$, the contribution from $f_{(2,3)}$ is constant over the entire feasible region. As such, we set $f_{(2,3)}(x_3) = 0$. Allowing g to be the standard Cantor function [3] on the interval $[0, 1]$, we define $f_{(2,1)}$ and $f_{(2,2)}$ by

$$f_{(2,1)}(x_1) = \begin{cases} 1 - x, & x_1 < 1 \\ 1, & x_1 \in [1, 2] \cap \mathbb{Q} \\ \sin(x - (4 + \pi)/2) + 1 & x_1 \geq 2 \\ 0, & \text{otherwise,} \end{cases}$$

and

$$f_{(2,2)}(x_2) = \begin{cases} \arctan(1 - x), & x_2 \leq 1 \\ 1 - 2g(x - 1), & 1 < x_2 < 4/3 \\ 2g(x - 1) - 1, & 4/3 \leq x_2 < 2 \\ \ln(x - 1), & x_2 \geq 2. \end{cases}$$

The geometry of these functions is depicted in Figure 2. Notice that the minimum values of both $F_1(x)$ and $F_2(x)$ are zero.

The function $f_{(2,1)}$ contributes the interval $[1, 2] \cup \{2 + k\pi : k = 0, 1, 2, \dots\}$ to H_1 . Since the only contribution from $f_{(1,1)}$ is zero, we have that

$$H_1 = \{0\} \cup [1, 2] \cup \{2 + k\pi : k = 0, 1, 2, \dots\}.$$

Let C be the set of Cantor ternary decimals, and set $C + 1 = \{y : y = w + 1 \text{ for some } w \in C\}$. With this notation we have that $f_{(2,2)}$ contributes $C + 1$ to H_2 . As before, $f_{(1,2)}$ contributes only zero to H_2 , which means that

$$H_2 = \{0\} \cup (C + 1).$$

Since $f_{(2,3)}$ is constant, H_3 only has contributions from $f_{(1,3)}$. Hence, $H_3 = \{0\}$.

Every feasible element has $x_3 = 0$, which means x_3 is always cornered. As a consequence, $\beta(x) \subseteq \{1, 2\}$ for every feasible x . Hence, $A_{\beta(x)}$ is a submatrix of $[0, 0]$, for every $x \in \mathcal{P}$. Obviously no nonempty subcollection of these columns can be independent, which means that every corner must have the property that $\beta(x) = \emptyset$ —i.e. every corner is degenerate. We conclude that the set of corners is $\mathcal{C} = H_1 \times H_2 \times H_3$.

Since $(1, 1, 0)^T$ is a corner, and $F((1, 1, 0)^T) = (0, 0)^T$, we have that there exists a corner optimal solution. However, $F((1, 3/2, 0)^T)$ also equals $(0, 0)^T$, but $(1, 3/2, 0)^T$ is not a corner. So, there are non-corner optimal solutions. These are depicted in Figure 2, where they are represented by all open line segment with $x_1 \in \mathbb{Q} \cap [1, 2]$, $x_3 = 0$, and x_2 ranging within $(1, 2)$. To see how the proof of Lemma 2 works, suppose that $x^0 = (1, 3/2, 0)^T$. Then, $\beta(x^0) = \{2\}$ and $A_{\beta(x^0)} = [0]$. We choose $q = (0, 1, 0)^T$, which means

$$\begin{aligned} \hat{\alpha}_+ &= \min \{5/3 - 3/2, 1 - 3/2\} = 1/6 \text{ and} \\ \hat{\alpha}_1 &= \min \emptyset = \infty. \end{aligned}$$

So, $\hat{\alpha} = 1/6$. The value of $\bar{\alpha}$ may be chosen to be any value in $(0, 1/6)$ (by an appropriate choice of ε_2), and we choose $\bar{\alpha} = 1/8$. Setting $x(\alpha) = x^0 + \alpha q$, we see that $F(x(\alpha)) = (0, 0)^T$, for all $\alpha \in (-1/8, 1/6)$, and hence is strongly monotonic over this line segment. Since $x^1 = x^0 + \hat{\alpha}q = (1, 5/3, 0)$ is a corner, we are done after one application of Lemma 2.

We conclude by showing that the Fundamental Theorem of Linear Programming follows from Theorem 2. We make the assumption that $(LP) \min\{c^T x : Ax = b, x \geq 0\}$ has an optimal solution. For $j = 1, 2, \dots, n$, define $d_j : \mathbb{R}^n \rightarrow \mathbb{R}$ by

$$d_j(x) = \begin{cases} -x_j, & x < 0 \\ 0, & x \geq 0, \end{cases}$$

and set $D(x) = \sum_{j=1}^n d_j(x)$. Notice that D has a minimum value of zero over \mathbb{R}_+^n . So, the definition of D , the fact that (LP) is feasible, and the lexicographic ordering imply that

$$\operatorname{argmin} \left\{ \begin{pmatrix} D(x) \\ c^T x \end{pmatrix} : Ax = b \right\} = \operatorname{argmin} \{c^T x : Ax = b, x \geq 0\}. \quad (1)$$

We denote the multiple objective program on the left-hand side of the equality in (1) by MOP' . Since the objective functions in MOP' are piecewise linear, they meet the conditions of Theorem 2, and hence MOP' has a corner optimal solution. Let x be a corner optimal solution to MOP' . Since the second objective in MOP' is linear, we have that $H_j = \{0\}$, for $j = 1, 2, \dots, n$. This means that x has the following properties: 1) $N(x) = \nu(x)$, 2) $x_N = x_\nu = 0$, and 3) the columns of $A_{B(x)} = A_{\beta(x)}$ are linearly independent. The third observation, together with the fact that x is optimal to (LP) from (1), implies that x is a basic optimal solution to (LP) . Hence, every corner optimal solution to MOP' is a basic optimal solution to (LP) , and Theorem 2 implies the Fundamental Theorem of Linear Programming.

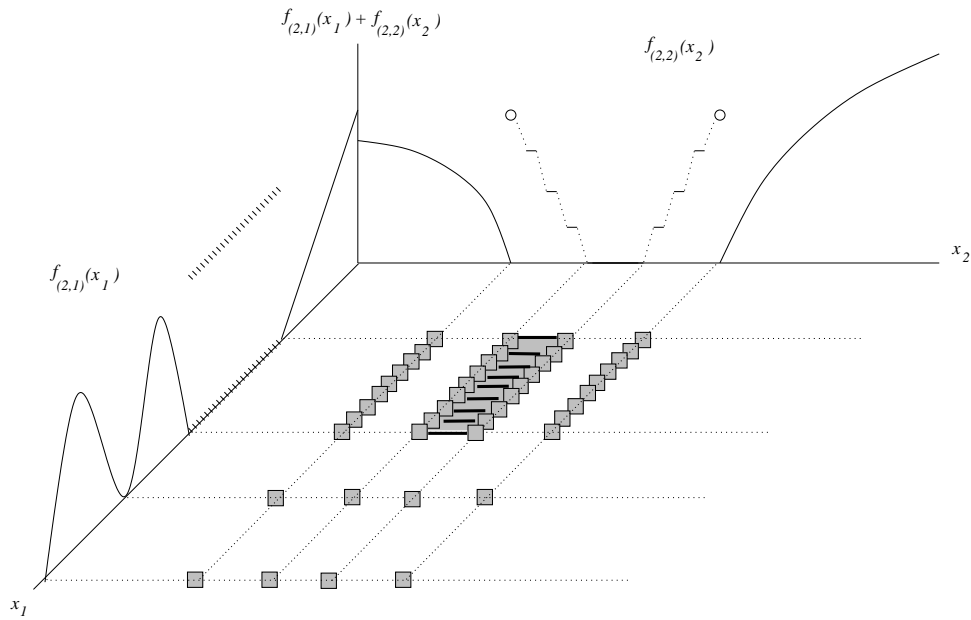


Figure 3: The geometry of $f_{(2,1)}(x_1)$ and $f_{(2,2)}(x_2)$ in Example 1. The shaded squares are optimal corners and dark lines between the shaded squares are non-corner optimal solutions.

3 Conclusion

We have provided a generalization of the Fundamental Theorem of Linear Programming, and our new result depends on index sets that use information from both the constraints and the objective function. This is in contrast to the linear programming case, where the index sets are defined independently of the objective function. The new index sets allow the concept of a basic solution to be extended to the concept of a corner. The objective functions allowed by Assumptions 1 and 2 may be quite complicated as demonstrated in Example 1. Moreover, Theorem 2 encompasses the situation of multiple objectives.

Acknowledgments

The authors would like to thank Harvey Greenberg, Ken Hummel, Stan Payne, and Jeff Stuart for comments on earlier revisions of this work.

References

- [1] Greenberg, H., *Mathematical Programming Glossary*, World Wide Web, <http://www.cudenver.edu/~hgreenbe/glossary/glossary.html>, 1996-2001.
- [2] Luenberger, D., *Linear and Nonlinear Programming*, 2nd ed., Addison Wesley, 1989.
- [3] Royden, H., *Real Analysis*, 3rd ed, Macmillan Publishing Company, 1988.
- [4] Ukovich, W., Pastore, S., and Premoli, A., *An Algorithm for Linearly Constrained Piecewise Lexicographic Programming Problems*, *Journal of Optimization Theory and Applications*, 111 (2001) 195-226.