To the Graduate Council:

I am submitting herewith a *thesis* written by Liet Anh Do entitled "*Duality in Linear Inequalities and Linear Programming*." I have examined the final copy of this *thesis* for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Master of Science, with a major in Mathematics.

Yueh-er Kuo

Dr. Yueh-Er Kuo, Major Professor

We have read this thesis
and recommend its acceptance:

Accepted for the Council:

Associate Vice Chancellor
and Dean of the Graduate School

STATEMENT OF PERMISSION TO USE

In presenting this thesis in partial fulfillment of the requirements for a Master's degree at the University of Tennessee, Knoxville, I agree that the Library shall make it available to borrowers under rules of the Library. Brief quotaions from this thesis are allowable without special permission, provided that accurate acknowledgement of the source is made.

Permission for extensive quotation from or reproduction of this thesis may be granted by my major professor, or in her absence, by the Head of Interlibrary Services when, in the opinion of either, in the proposed use of the material is for scholarly purposes. Any copying or use of the material in this thesis for financial gain shall not be allowed without my written permission.

Signature ______________________

Date ______________________

DUALITY IN
LINEAR INEQUALITIES AND LINEAR PROGRAMMING

A Thesis
Presented for the
Master of Science
Degree
The University of Tennessee, Knoxville

Liet Anh Do

August 1994

ACKNOWLEDGEMENTS

The author wishes to express sincere gratitude to Dr. Yueh-er Kuo for her suggestions and aid in preparing this paper. Appreciation is also extended to Dr. K. R. Stephenson and Dr. W. R. Wade for part of knowledge over the past years, especially for their time reading this thesis.

ABSTRACT

The subject of this paper is the theory of linear inequalities and linear programming. It aims, through use of theory of the duals, to set some formal connections between linear inequalities and linear programming problems, as well as to lay theoretical foundations for certain practical algorithms having been used to solve a general linear program.

In particular, this paper will attempt to present, at the end, the simplex algorithm invented by Dantzig, and the primal-dual algorithm developed later by Dantzig, Ford, and Fulkerson in terms of the general theory, and the duality theory of linear programs. Once the problems of foundation are uncovered, certain features of both algorithms will be presented, from which we could evaluate which one would be more advantegeous than the other.

TABLE OF CONTENTS

CHAPTER I

INTRODUCTION

Since linear inequalities are part of a linear program, this paper will develop chapters for linear inequalities and linear programs separately in such a way that properties established for linear inequalities will be used to derive properties of linear programs. Moreover, duality will be developed for linear inequalities and linear programs in additional chapters to indicate how the concept of the duals are used to reduce a linear programming problem to a linear inequality problem, as well as to develop different algorithms.

The presentation of the paper is divided into four chapters presented primarily in matrix notation. Chapter II begins with some basic definitions related to the systems of linear inequalities, emphasizing on the geometric ideas accompanying their solution sets. The theory of faces will be discussed and employed to prove Minkowski's theorem [6] in that a nice structure of the solution set is presented. Finally, the resolution theorem will be covered to give more structural ideas about the solution sets.

In chapter III, the dual systems of linear inequalities will be discussed. Basic definitions, constructive lemmas are developed in details to make way for Farcas' theorem [6] whereby a structure for a "dual" solution set is presented. Finally, the separation theorem will be examined to give more geometric insight into a pair of special solution sets.

As the foregoing chapters focus on the system of linear inequalities, the last two chapters will be devoted to linear programming problems, benifiting at hand the consequences developed for the system of inequalities as a useful knowledge for the system of linear constraints in a linear program. Chapter IV is mainly concerned with general linear programs. Extreme points and extreme directions will be presented as main concepts from which a feasible solution and an optimal solution of a linear program will be respectively characterized. The simplex method will be discussed next, based on general ideas of Theorem 4.8. And finally, the tableau simplex algorithm will be formulated, using ideas elaborated in Theorem 4.10, to fecilitate computations.

For the last chapter, chapter V, the dual linear programs will be discussed. Many concepts and formal properties are mentioned; especially, the duality theorem and the existence theorem will be illustrated and utilized to establish Theorem 5.6 in which the concept of complimentary slackness is presented. This concept will be used as a main source to improve the two-phase simplex algorithm, also presented in this chapter, into the primal-dual algorithm, which will be discussed in great detail for the remainder of the chapter.

CHAPTER II

LINEAR INEQUALITIES

We will begin the chapter by giving some basic definitions related to linear inequalities, and then develop Minkowski's Theorem via using the theory of faces. Finally, the resolution theorem will be disscussed. The main references of this chapter are Gale [3], Goldman [5], and Goldman and Tucker [6].

Unless otherwise specified, E_n will denote the Euclidean n-dimensional real space. Using 'T' for matrix transpose, we first state a definition.

DEFINITION 2.1. A function $L : E_n \to E_1$, defined by

$$L(X) = C^T X, \tag{2.1}$$

where

$X = [x_1\ x_2\ ..\ x_n]^T$, and $C = [c_1\ c_2\ ..\ c_n]^T$ (fixed in E_n),

is called a **Linear Function**.

DEFINITION 2.2. Let b_i in E_1 $(i = 1,2,..,m)$. If there exists some vector X in E_n such that

$$L_i(X) \leq b_i, \quad i = 1,2, . . ,m, \tag{2.2}$$

where L_i's are linear functions, then we say that (2.2) is a **Consistent System of Linear Inequalities**. Otherwise, we say that it is an **Inconsistent System of Linear Inequalities**.

A consistent or inconsistent system of linear inequalities will be called a **System of Linear Inequalities**, and hence the form (2.2) can be named regardless of its consistency.

DEFINITION 2.3. The set defined by

$$\{X \mid L_i(X) \leq b_i,\ i = 1,2,\ .\ .\ ,m\} \tag{2.3}$$

will be called the **Solution Set** the system (2.2).

Clearly, (2.3) is nonempty if and only if the system (2.2) is consistent.

By letting

$$B = [b_1\ b_2\ ..\ b_m]^T,$$

and

$$A = [a_{ij}],\ j = 1,2,.\ .,n,$$

where

$$\Sigma_{1 \leq j \leq n}\ a_{ij}x_j = L_i(X),\ i = 1,2,..,m,$$

we put the system of linear inequalities (2.3) in a matrix form:

$$AX \leq B. \tag{2.4}$$

When B is a vector of zeros, the system will be called **Homogeneous**. Otherwise, it will be called **Nonhomogeneous**.

Before going further into part of the theory related to dual linear inequalities, we develop some geometric concepts and properties related to general linear inequalities. We first recognize geometric ideas, associated with the solution set of a system of linear inequalities, in the following definitions, noticing the fact that these geometric ideas will motivate the use of language.

DEFINITION 2.4. Given a vector $[a_j]^T$ in E_n, and a number b in E_1. Then, the set

$$\{c[a_j]^T \mid c \geq 0\}$$

will be called a **Half Line**, and the set

$$\{[x_j]^T \in E_n \mid \Sigma_{1\leq j\leq n} a_j x_j = b\}$$

will be called a **Hyperplane**.

A hyperplane geometrically divides E_n into two **Half Spaces**, formally defined by the sets

$$\{[x_j]^T \in E_n \mid \Sigma_{1\leq j\leq n} a_j x_j \leq b\},$$

and

$$\{[x_j]^T \in E_n \mid \Sigma_{1\leq j\leq n} a_j x_j \geq b\}.$$

The intersection of finitely many halfspaces will be called a **Polyhedral Convex Set**, which is, in fact, the solution set of the system (2.4). The solution set of a homogeneous system of linear inequalities will be called a **Polyhedral Convex Cone**. We also call a polyhedral convex set **Vacuous** when the system (2.2) is not consistent, and call a polyhedral convex cone **Trivial** when the homogeneous system has the unique solution, a vector of zeros.

From now on, unless differently stated, A will denote an mxn real matrix, and A_j, j = 1,2,..,n, will denote a Column

Vector of A, except for this chapter, where it means a row vector, i.e., a one-column matrix whose entries are identified with those in the corresponding row of A. Also, **I** will denote an identity matrix, **O** will denote a matrix of 0's, **1** will denote a one-column matrix of 1's, and $\mathbf{e}_i$ will denote a one-column matrix of 0's except for 1 at the i-th position, having the matrix sizes determined by the context. As a matter of convenience, we sometimes write x_j or $(X)_j$ for the j-component of a vector X, a_{ij} or $(A)_{ij}$ for the ij-component of a matrix A, and $(x_1,x_2,..,x_n)$ for $[x_1\ x_2\ ..\ x_n]^T$.

The following definition will be used to construct Lemma 2.1, and Theorem 2.1, from which Minkowski's Theorem [6] is derived.

DEFINITION 2.5. Let A* be defined by

$$A^* = \{X|\ AX \le \mathbf{O}\}, \tag{2.5}$$

and

$$H \subseteq I = \{1,2,..m\}.$$

Then, a **Face** of A*, corresponding to H, is defined by

$$F_H = \{X|\ A_h^T X < 0 \text{ for each h in H}\} \cap L_H,$$

where

$$L_H = \{X|\ A_h^T X = 0 \text{ for each } h \in I/H\}.$$

The **Dimension** of a face F_H is the dimension of L_H, i.e.,

$$d_H = n - r_H,$$

where r_H is the maximal number of linearly independent vectors in the set $\{A_h | \ h \in I/H\}$.

If A is of rank n, then $d_\emptyset = 0$ and $F_\emptyset$ becomes the unique zero-dimensional face, called the **Vertex** of A*, and the resulting one-dimensional faces will be called the **Edges** of A*.

If G is a proper subset of H such that both F_H and F_G are both nonempty, then F_G is called a **Boundary Face** of F_H. The union of all boundary face of F_H will be called the **Boundary** of F_H.

EXAMPLE 2.1. Consider A* with

$$A = \begin{bmatrix} -1 & 1 \\ 1 & -2 \end{bmatrix}.$$

Find all of its faces and edges.

SOLUTION. Geometrically, A* can be easily pictured in E_2. The nonvacuous faces are

$$F_{12} = \{(x_1, x_2) \mid x_2 < x_1\} \cap \{(x_1, x_2) \mid 2x_2 > x_1\} \qquad (d_{12} = 2),$$

$$F_1 = \{(x_1, x_2) \mid x_2 < x_1\} \cap \{(x_1, x_2) \mid 2x_2 = x_1\} \qquad (d_1 = 1),$$
$$= \{(x_1, x_2) > \mathbf{0} \mid 2x_2 = x_1\}$$

$$F_2 = \{(x_1, x_2) \mid x_2 = x_1\} \cap \{(x_1, x_2) \mid 2x_2 > x_1\} \qquad (d_2 = 1),$$
$$= \{(x_1, x_2) > \mathbf{0} \mid x_2 = x_1\}$$

$$F_\emptyset = \{(x_1, x_2) \mid x_2 = x_1\} \cap \{(x_1, x_2) \mid 2x_2 = x_1\} \qquad (d_\emptyset = 0),$$
$$= \{\mathbf{0}\}.$$

The edges of A* are F_1, F_2, which can also have the forms:

$$F_1 = \{tX_1 \mid t > 0\}, \text{ where } X_1 = (1, 1/2);$$

$$F_2 = \{tX_1 |\ t > 0\},\text{ where } X_2 = (1,1).$$

LEMMA 2.1. For each vector X_0 in a face F_H that has dimension $d_H > d_\emptyset + 1$, we have

$$X_0 = X_1 + X_2,$$

where X_1, X_2 are in some boundary faces (of F_H) of dimension $> d_\emptyset$ but $< d_H$.

PROOF. Let L_H be a linear space of dimension d_H. Then, $F_H \subseteq L_H$ and there exists a vector X' in L_H such that X' is not in the linear subspace L generated by X_0 and $F_\emptyset$ (since $d_H > d_\emptyset + 1$). Thus, X_0 and X' span a 2-dimensional subspace L_0 of L_H that consists of vectors X of the form:

$$X = tX_0 + sX'. \tag{2.7}$$

Consider

$$M_0 = L_0 \cap A^*.$$

Then,

$$X \text{ is in } M_0$$

if and only if

$$A_h^T X = t(A_h^T X_0) + s(A_h^T X') \leq 0 \text{ for each } h. \tag{2.8}$$

Note that for each $h \notin H$, the form (2.7) satisfies the form (2.8). This implies that

$$X \in M_0$$

if and only if

$$A_h^T X = t(A_h^T X_0) + s(A_h^T X') \leq 0 \text{ for each } h \text{ in } H.$$

Since

$$d_H > d_\emptyset + 1 > d_\emptyset,$$

we have

$$H \neq \emptyset.$$

To determine s and t, we consider the graphs of the equations

$$t(A_h^T X_0) + s(A_h^T X') = 0, \ h \in H,$$

in the (s,t) plane, which are the straight lines going through the origin with the slopes

$$f_h = -\frac{A_h^T X'}{A_h^T X_0}, \ h \in H.$$

Note that

$$A_h^T X_0 < 0 \text{ for each } h \text{ in } H.$$

So, (2.8) can be reduced to the form

$$t \geq f_h s. \tag{2.9}$$

It can be seen that the set of all straight lines must have at least two elements. Hence, by letting $f_{h'}$, $f_{h''}$ be the maximal and the minimal slopes, corresponding to some h', and h" in H, we can define

$$s_1 = \frac{1}{f_{h'} - f_{h''}}, \ t_1 = \frac{f_{h'}}{f_{h'} - f_{h''}},$$

and

$$s_2 = -\frac{1}{f_{h'} - f_{h''}}, \ t_2 = -\frac{f_{h''}}{f_{h'} - f_{h''}}.$$

Thus, we obtain

$$X_1 = t_1X_0 + s_1X', \ X_2 = t_2X_0 + s_2X',$$

so that

$$X_1 + X_2 = X_0.$$

Moreover, (2.9) can be reduced to

$$f_{h'} \geq f_h \geq f_{h''}, \ \forall \ X_1, X_2.$$

Now, X_1 satisfies (2.9), and hence (2.8), as as an equation for h = h', and as a strict inequalitiy for h = h", while X_2 does the same thing respectively for h = h", and h = h'. And thus, each of X_1 and X_2 satisfies (2.8) as a strict inequality for some, but not all, h in H.

Since X_1, and X_2 are in L_H, they satisfy (2.8) as an equation for each h $\notin$ H. Hence, each of X_1 and X_2 must lie in a boundary face of F_H that has dimension > $d_\emptyset$. ■

EXAMPLE 2.2. Illustrate Lemma 2.1, using Example 2.1, with a particular vextor X_0 = (3/2,1).

SOLUTION. We may optionally choose X' = (-1,1). Then,

$$f_1 = \frac{(-1,1)(-1,1)}{(-1,1)(\frac{3}{2},1)} = 4,$$

$$f_2 = \frac{(1,-2)(-1,1)}{(1,-2)(\frac{3}{2},1)} = -6,$$

$$s_1 = \frac{1}{4+6} = \frac{1}{10}, \quad s_2 = \frac{-1}{4+6} = -\frac{1}{10},$$

and

$$t_1 = \frac{4}{4+6} = \frac{2}{5}, \quad t_2 = \frac{6}{4+6} = \frac{3}{5}.$$

It follows that

$$X_2 = t_2 X_0 + s_2 X' = \frac{3}{5}(\frac{3}{2},1) - \frac{1}{10}(-1,1) = (1,\frac{1}{2}),$$

$$X_1 = t_1 X_0 + s_1 X' = \frac{2}{5}(\frac{3}{2},1) + \frac{1}{10}(-1,1) = (\frac{1}{2},\frac{1}{2}),$$

and

$$X_1 + X_2 = (\frac{1}{2},\frac{1}{2}) + (1,\frac{1}{2}) = (\frac{3}{2},1) = X_0.$$

It should be noted that the face $F_\emptyset$ can be identified with the null space of A. Since the null space of A and its dimension are unique, so are $F_\emptyset$ and $d_\emptyset$. Now, suppose that there exists a face F_H of dimension $< d_\emptyset$. Then, we have

$$F_H = \{X|\ A_h^T X < 0,\ \forall\ h \in H\} \cap \{X|\ A_h^T X = 0,\ \forall\ h \notin H\},$$

and

$$\text{dimension of } F_H = \text{dimension of } \{X|\ A_h^T X = 0,\ \forall\ h \notin H\}$$
$$\geq \text{dimesion of } \{X|\ A_h^T X = 0,\ \forall\ h \in I\} = d_\emptyset.$$

Therefore,

$$d_\emptyset > d_\emptyset,$$

which is impossible. It follows that there exists no face of dimension $< d_\emptyset$.

DEFINITION 2.6. The set A^*, defined in Definition 2.5, has the following properties, which are called **Nonnegative Linear**:

$$\text{for } k \geq 0, \text{ and } X \in A^*, \; kX \in A^*, \tag{2.10}$$

$$\text{for } X_1, X_2 \in A^*, \; X_1 + X_2 \in A^*. \tag{2.11}$$

These properties are equivalent to the following one:

$$\text{For } k_1, k_2 \geq 0, \; X_1, X_2 \in A^*, \quad k_1X_1 + k_2X_2 \in A^*,$$

whereby the sum $k_1X_1 + k_2X_2$ is called a **Nonnegative Linear Combination** of X_1 and X_2, which can be defined more extensively for any finite number of vectors.

A set $C \subseteq E_n$ will be called a **Convex Cone** if it is nonnegative linear. A set $C \subseteq E_n$, containing a set B, will be called a **Convex Cone Hull** of B, if each element of C can be express as a nonnegative linear combination of elements in B. When this is the case, we also say that B **Spans** C, and write $B^{\angle} = C$. In the same context, if each element of C can be written as a **Finite Convex Combination** of elements in B, i.e., for each X in C,

$$X = \sum_{j=1}^{n} k_j B_j, \; \text{where } k_j \geq 0, \; B_j \in B, \; \sum_{j=1}^{n} k_j = 1,$$

then C will be called a **Convex Hull** of B. And when the context is clear, we also say that B **Spans** C.

Clearly, if B spans C, then C is the smallest convex cone or the smallest convex set that contains B. In addition, if B is finite, then C is a polyhedral convex cone or a bounded convex polyhedron defined later in Definition 2.8. In particular, Let $B = \{B_1, B_2, .., B_n\}$ be a finite set of vectors in E_m. Then,

$$B^{\angle} = \{\textstyle\sum_{1\le j\le n} v_j B_j \mid v_j \ge 0,\ j = 1,2,..,n\}. \qquad (2.12)$$

Identifying B as a matrix whose columns are B_j's, (2.12) can be expressed in the form:

$$B^{\angle} = \{BV \mid V \ge \mathbf{0}\}.$$

Moreover, $B^{\angle}$ is a convex hull of half lines

$$\{v_j B_j \mid v_j \ge 0\},\ j = 1,2,..,n,$$

generated by B_j's in B.

On the other hand, identify A with $\{A_1, A_2, .., A_m\}$, where $A_1, A_2, ..,$ and A_m are the row vectors of A. Then, we can put A* in an equivalent form:

$$A^* = \{X \mid A_1^T X \le 0,\ A_2^T X \le 0, .., A_m^T X \le 0\}.$$

This new characterization of A* can be generalized into the following definition for an arbitrary subset S of E_n.

DEFINITION 2.7. (**Polar**) Let $S \subseteq E_n$. Then, the set

$$\{X \in E_n \mid X^T Y \le 0 \text{ for all } Y \text{ in } S\}$$

is called the **Polar** of S, and denoted by S*. Also, S** will be denoted for (S*)*.

It is in preceding general senses that an mxn matrix M will have the following properties:

$$M^{\angle\angle} = M^{\angle}, \qquad (2.13)$$

$$M^{***} = M^*, \qquad (2.14)$$

$$((M^{\top})^*)^{\angle} = (M^{\top})^* = (M^{\angle})^*. \tag{2.15}$$

THEOREM 2.1. A^* is either exactly the (unique) $d_\emptyset$-face, or the convex hull of the $d_\emptyset$-face, and the $(d_\emptyset+1)$-faces.

PROOF. As mentioned in the statement preceding Definition 2.6, A^* has no face of dimension $< d_\emptyset$. So, suppose that A^* has more than one $d_\emptyset$-face. Then, for the vectors in the $(d_\emptyset+1)$-faces, they must automatically lie in the convex hull of these faces. As for any vector X in a face of dimension $> (d_\emptyset+1)$, we can use Lemma 2.1 repeatedly to have X as a finite sum of vectors each lying in a face of dimension $d_\emptyset+1$, i.e.,

$$X = X_1 + X_2 + .. + X_m.$$

Hence,

$$X = (1/m)(mX_1) + (1/m)(mX_2) + .. + (1/m)(mX_m).$$

Since each mX_i, $i = 1,2,..,m$, is in the same $(d_\emptyset+1)$-face where X_i lies, X is also in the convex hull C of these $(d_\emptyset+1)$-faces. Note that A^* is a disjoint union of all the faces. Hence, $A^* \subseteq C$. In addition, since A^* is convex (and certainly contains the faces of dimension $d_\emptyset+1$), it must contain C. ■

Before giving corollaries which are direct results of Theorem 2.1, it should be noted that $F_\emptyset = L_\emptyset$, and thus it is a linear subspace of E_n.

COROLLARY 1.1. The intersection of finitely many half spaces is either a linear subspace of dimension $d_\emptyset$ or the convex hull of finitely many half subspaces of dimension $d_\emptyset+1$. (Here, **Dimension of a Half Space** is defined to be the dimension of the linear space that it spans. And when B = **O** (related to Definition 2.3), the name **Half Subspace** will be used for half space.)

COROLLARY 1.2. If the rank of A is n, then A* is either $\{\mathbf{O}\}$ or the convex hull of $\{\mathbf{O}\}$ and the edges of A*.

We next present a theorem that gives a characterization for a convex polyhedral cone.

THEOREM 2.2. (**Minkowski's Theorem**) For each polyhedral convex cone A*, there exists a finite set B = $\{B_1,B_2,..,B_q\}$ of vectors in E_n such that A* = $B^{\angle}$.

PROOF. we need to construct a set B that spans A*. Using Theorem 2.1, it suffices to construct B as a union of finite sets which spans the face $F_\emptyset$ and each of the $(d_\emptyset+1)$-faces, one by one.

If $d_\emptyset = 0$, then take $S_\emptyset = \{\mathbf{O}\}$ as a spanning set of $F_\emptyset$. Otherwise, with $d_\emptyset > 0$, take $\{B_1,B_2,..,B_d\}$, $d = d_\emptyset$, equal to a basis of $F_\emptyset$. Then, for each X in $F_\emptyset$, we can write

$$X = \sum_{i=1}^{d_\emptyset} c_i B_i .$$

Thus, for each i, set

$$u_i = \max \{c_i, 0\} \geq 0,$$

and

$$v_i = \max \{-c_i, 0\} \geq 0.$$

Then,

$$X = \sum_{i=1}^{d_\emptyset} u_i B_i + \sum_{i=1}^{d_\emptyset} -v_i B_i .$$

This implies that

$$S_\emptyset = \{B_i\}_{i=1}^{d_\emptyset} \cup \{-B_i\}_{i=1}^{d_\emptyset}$$

is a spanning set of $F_\emptyset$.

We next consider an arbitrary $(d_\emptyset+1)$-face F_H. Then, write

$$F_H = O_H \cap L_H,$$

where

$$O_H = \{X | \ A_h^T X < 0 \text{ for each h in H}\},$$

$$L_H = \{X | \ A_h^T X = 0 \text{ for each h in I/H}\}.$$

It can be seen that L_H is a $(d_\emptyset+1)$-dimensional linear space that contains $F_\emptyset$. So, let $B_H \in F_H \subset L_H$, then $B_H \in L_H/F_\emptyset$.

Furthermore, for each vector X in F_H, we can write

$$X = X_\emptyset + cB_H \text{ for some } X_\emptyset \text{ in } F_\emptyset.$$

Multiplying both sides by A_h for some h in H, we have

$$A_h^T X = A_h^T X_\emptyset + cA_h^T B_H,$$

where

$$A_h^T X > 0,\ A_h^T X_\emptyset = 0,\ \text{and } A_h^T B_H > 0.$$

Hence,

$$c > 0,$$

and

$$S_H = S_\emptyset \cup \{B_H\}$$

is a spanning set for F_H. ∎

EXAMPLE 2.3 Consider the polyhedral convex cone A^*, where

$$A = \begin{bmatrix} -1 & 1 \\ 1 & -2 \end{bmatrix}.$$

Then, construct B such that $A^* = B^\angle$.

SOLUTION. As shown in Example 2.1, $d_\emptyset = 0$. Hence, take $S_\emptyset = \{(0,0)\}$ as a spanning set of $F_\emptyset$. Next, we consider the $(d_\emptyset+1)$-faces, which were obtained in the previous example:

$$F_1 = \{(x_1,x_2) \mid x_1 = 2x_2\} = O_1 \cap L_1,$$

where

$$O_1 = \{(x_1,x_2) \mid x_1 > x_2\},$$
$$L_1 = \{(x_1,x_2) \mid x_1 = 2x_2\}.$$

Note that L_1 is a one-dimensional space and contains $F_\emptyset = \{(0,0)\}$. So, we can optionally pick $B_1 = \{(2,1)\}$. Similarly, we pick $B_2 = \{(1,1)\}$, associated with F_2. As a result, we have

$$B = \{(0,0),(2,1),(1,1)\},\ \text{and } A^* = B^\angle.$$

EXAMPLE 2.4. To illustrate Theorem 2.2, with the case $d_\emptyset > 0$, we consider the polyhedral convex cone A*, where

$$A = \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix}.$$

Then,

$$F_\emptyset = \{(x_1,x_2) \mid x_1 = x_2\},$$

which is a one-dimensional space, and $d_\emptyset = 1$. Therefore, we can optionally take

$$B_1 = (1,1), \text{ and } B_2 = (-1,-1)$$

to form a spanning set for $F_\emptyset$. Note that

$$F_1 = \{(x_1,x_2) \mid x_1 > x_2\} \cap \{(x_1,x_2) \mid x_1 = x_2\} = \emptyset,\ d_1 = 0,$$

$$F_2 = \{(x_1,x_2) \mid x_1 < x_2\} \cap \{(x_1,x_2) \mid x_1 = x_2\} = \emptyset,\ d_2 = 0,$$

$$F_{12} = \{(x_1,x_2) \mid x_1 < x_2 \text{ and } x_1 > x_2\} \cap E_2 = \emptyset,\ d_{12} = 0.$$

Hence, there is no face other than $F_\emptyset$. Consequently, we have

$$B = \{(1,1),(-1,-1)\}, \text{ and } A^* = B^\angle.$$

Besides, to display a situation with more than one face, we may consider A*, where

$$A = \begin{bmatrix} -1 & 1 \\ 0 & 0 \end{bmatrix}.$$

Then,

$$F_\emptyset = \{(x_1,x_2) \mid x_1 = x_2\},\ d_\emptyset = 1,\ (L_\emptyset = F_\emptyset),$$

$$F_1 = \{(x_1,x_2) \mid x_1 > x_2\},\ d_1 = 2,\ (L_1 = E_2),$$

$$F_2 = \{(x_1,x_2) \mid x_1 = x_2\},\ d_2 = 1,$$

$$F_{12} = \{(x_1,x_2) \mid x_1 = x_2\},\ d_{12} = 0.$$

Hence, we can take

$(1,1),(-1,-1)$ from $L_\emptyset$, and $(2,1)$ from F_1

to have

$$B = \{(1,1),(-1,-1),(2,1)\}, \text{ and } A^* = B^\angle.$$

EXAMPLE 2.5. Consider

$$A = \begin{bmatrix} -3 & -2 & 6 \\ -1 & 1 & -1 \\ 1 & -2 & -2 \\ 1 & -8 & 4 \end{bmatrix}.$$

Then, construct B such that $B^\angle = A^*$.

SOLUTION. Note that since $d_\emptyset = 0$, only one-dimensional faces are used to compute the spanning vectors. Optionally, we pick

$$(4,1,1) \in F_{12},\ (4,3,-1) \in F_{14},$$
$$(20,9,13) \in F_{23},\ (4,9,5) \in F_{34}.$$

Hence, we obtain

$$B = \{(4,1,1),(4,3,-1),(20,9,13),(4,9,5)\}, \text{ and } A^* = B^\angle.$$

The above theorem immediately implies the following corollary.

COROLlARY 2.2.1. The intersection of finitely many half spaces is a convex hull of finitely many half lines.

The next purpose is to see how the solution set of a nonhomogeneous system of linear inequalities can be broken down and expressed in terms of a solution set of an homogeneous system of linear inequalities. As this is done, we have a constructive way to characterize all solutions of a nonhomogeneous system by means of a finite number of vectors, which are also the solutions of the system and called the **Fundamental Set of Solutions**. For this purpose, we first need a lemma that associates a polyhedral convex set S in E_n to a polyhedral convex cone in which S is embedded.

LEMMA 2.2. The rule

$$S = \{X \mid AX \le B\} \rightarrow C_{n+1} = \{(X,t) \mid AX - Bt \le \mathbf{0},\ t \ge 0\}$$

is inclusion-preserving, well-defined, one-to-one, and maps any nonvacuous polyhedral convex set S to a polyhedral convex cone that lies in the half space $\{(X,t) \in E_{n+1} \mid t \ge 0\}$ and meets the hyperplane $\{(X,t) \in E_{n+1} \mid t = 1\}$. (Hence, it defines a correspondence $C_{n+1}(S)$ between the collection of all nonvacuous polyhedral convex sets $S \subseteq E_n$ and its range consisting of polyhedral convex cones $C_{n+1} \subseteq E_{n+1}$, so that we can also write $S(C_{n+1})$ as the inverse map of $C_{n+1}(S)$.)

PROOF. a) Let

$$S = \{X \mid AX \le B\}$$

be a nonvacuous subset of

$$T = \{X \mid CX \le D\},$$

and

$$X_0' = (X_0, t_0) \in \{(X,t) \mid AX - Bt \le \mathbf{0},\ t \ge 0\}.$$

If

$$t_0 > 0,$$

then

$$X_0/t_0 \in S \subseteq T,$$

so that

$$X_0' \in \{(X,t) \mid CX - Dt \leq \mathbf{0},\ t \geq 0\}.$$

If

$$t_0 = 0,$$

then

$$X_0 \in \{X \mid AX \leq \mathbf{0}\}.$$

Fix $X_1 \in S \subseteq T$. Then,

$$AX_1 \leq B, \text{ and } AX_0 \leq \mathbf{0}.$$

It follows that

$$AX_1 + \alpha AX_0 = A(X_1 + \alpha X_0) \leq B,$$

and hence

$$X_1 + \alpha X_0 \in S \subseteq T \text{ for any } \alpha \geq 0.$$

Therefore,

$$CX_1 + \alpha CX_0 \leq D. \tag{2.16}$$

Suppose that $CX_0 \nleq \mathbf{0}$. Then, By denoting the i-th component of a vector V by $(V)_i$, we have

$$(CX_0)_i > 0 \text{ for some } i,$$

so that

$$\alpha(CX_0)_i > (D)_i - (CX_1)_i \text{ for some } \alpha \geq 0.$$

But this contradicts (2.16). It follows that

$$CX_0 \leq \mathbf{0},$$

so that

$$X_0 \in \{X \mid CX \leq \mathbf{0}\}.$$

Therefore,

$$(X_0, 0) \in \{(X,t) \mid CX - Dt \leq \mathbf{0},\ t \geq 0\}.$$

This implies that

$$\{(X,t) \mid AX - Bt \leq \mathbf{0},\ t \geq 0\} = \{(X,t) \mid CX - Dt \leq \mathbf{0},\ t \geq 0\},$$

and hence the rule is inclusion-preserving.

b) To show that the rule is well-defined, suppose that $S = T$. Then, from (a), it immediately follows that

$$\{(X,t) \mid AX - Bt \leq \mathbf{0},\ t \geq 0\} \subseteq \{(X,t) \mid CX - Dt \leq \mathbf{0},\ t \geq 0\},$$

so that the rule is well-defined.

c) To show that it is one-to-one, first notice that

$$S = \{X \mid (X,1) \in C_{n+1}\}. \tag{2.17}$$

Suppose that S, T are mapped to C_{n+1}, and D_{n+1}, respectively, such that $C_{n+1} = D_{n+1}$. Then, by (2.17), we have that

$$S = \{X \mid (X,1) \in C_{n+1}\} = \{X \mid (X,1) \in D_{n+1}\} = T.$$

Therefore, the rule is one-to-one.

d) To prove the remaining part, suppose that $S \neq \emptyset$. Then, there exists a vector X_0 in S such that $AX_0 \leq B$, i.e.,

$$AX_0 - B.1 \leq \mathbf{0}.$$

It follows that

$$(X_0, 1) \in C_{n+1},$$

and hence, C_{n+1} lies in the half space $t \geq 0$ and meets the half plane $t = 1$. ■

It should be noted that the inverse map $S(C_{n+1})$ of Lemma 2.2 will map any polyhedral convex cone D_{n+1} that lies in the half space $t \geq 0$, and meets the hyperplane $t = 1$, to a nonvacuous polyhedral convex set. To see this, write

$$D_{n+1} = \{(X,t) \mid AX - Bt \leq \mathbf{0},\ t \geq 0\},$$

and note that since D_{n+1} meets the hyperplane $t = 1$, there exists $(X_0,1) \in D_{n+1}$ such that $AX_0 \leq B$. This implies $X_0 \in S$, i.e., $S \neq \emptyset$.

DEFINITION 2.8. Let $P = \{P_1, P_2, .., P_n\}$ be the set of vectors in E_m. Then, a set

$$P^{\Delta} = \{\sum_{j=1}^{n} u_j P_j \mid u_j \geq 0,\ \sum_{j=1}^{n} u_j = 1\}$$

will be called a **Bounded Convex Polyhedron.**

Identifying P as a matrix whose columns are P_j's, we can put the above expression in an equivalent form:

$$P^{\Delta} = \{PU \mid U \geq \mathbf{0}, \sum_{j=1}^{n} u_j = 1\}.$$

Clearly, P^{Δ} is presisely a convex hull of a finite number of vectors, namely $P_1, P_2, .., P_n$. Moreover, it can be used to construct the following theorem.

THEOREM 2.3. **(Resolution Theorem)** For each nonvacuous polyhedral convex set $S = \{X \mid AX \leq B\}$, there exists a bounded convex polyhedron

$$P^{\Delta} = \{PU \mid U \geq \mathbf{0}, \sum_{j=1}^{n} u_j = 1\},$$

and a polyhedral convex cone

$$Q^{\angle} = \{QV \mid V \geq \mathbf{0}\}$$

such that

$$S = P^{\Delta} + Q^{\angle}.$$

Conversely, each nonvacuous set of the form $P^{\Delta} + Q^{\angle}$ is a polyhedral convex set.

PROOF. By Lemma 2.2, there exists a polyhedral convex cone $C_{n+1}(S)$ in the half space $t > 0$. So, by Corollary 2.2.1, there exists a finite set of vectors in E_{n+1}, which spans $C_{n+1}(S)$. This set can be written as the disjoint union of two sets:

$$\{(P_1,t_1), (P_2,t_2), .., (P_p,t_p)\}, \text{ where } t_i\text{'s} > 0,$$

and

$$\{(Q_1,0), (Q_2,0), .., (Q_p,0)\}.$$

Furthermore, since $C_{n+1}(S)$ is a cone, we may assume t_i's $= 1$ without loss of generality. Note that since $C_{n+1}(S)$ meets the hyperplane $t = 1$, we have $p > 0$. Besides, when $q = 0$, we set $(Q_1,0) = (0,0)$. From

$$S = \{X \mid (X,1) \in C_{n+1}(S)\},$$

it follows that

$$S = \{\sum_{i=1}^{p} u_i P_i + \sum_{j=1}^{q} v_j Q_j,\ u_i \geq 0,\ v_j \geq 0,\ \sum_{i=1}^{p} u_i = 1\}.$$

That is,

$$S = P^{\Delta} + Q^{\angle},$$

where

$$P = [P_1, P_2, \ldots, P_p], \text{ and } Q = [Q_1, Q_2, \ldots, Q_q].$$

Conversely, beginning with a nonvacuous set of the form

$$S = P^{\Delta} + Q^{\angle},$$

we first notice that

$$Q^{\angle} = \{X \mid AX \leq \mathbf{0}\}.$$

To see this, define

$$D_{n+1} = \{\sum_{i=1}^{p} u_i (P_i, 1) + \sum_{j=1}^{q} v_j (Q_j, 1)\}.$$

Then,

$$S = \{X \mid (X, 1) \in D_{n+1}(S)\}.$$

This implies that

$$S = S(D_{n+1}).$$

It follows from Lemma 2.2 that

$$D_{n+1} = C_{n+1}(S),$$

so that

$$\begin{aligned} Q^{\angle} &= \{X \mid (X, 0) \in D_{n+1}(S)\} \\ &= \{X \mid (X, 0) \in C_{n+1}(S)\} \\ &= \{X \mid AX \leq \mathbf{0}\}. \end{aligned}$$

Therefore,

$$P^{\Delta} + Q^{\angle} = \{X \mid (X,1) \in D_{n+1}(S)\},$$

where D_{n+1} lies in the half space $t = 1$. By Lemma 2.2,

$$S(D_{n+1}) = \{X \mid (X,1) \in D_{n+1}\}$$

is a nonvacuous polyhedral convex set. ■

EXAMPLE 2.6. Consider the set $S = \{X \mid AX \le E\}$, where

$$A = \begin{bmatrix} -3 & -2 \\ -1 & 1 \\ 1 & -2 \\ 1 & -8 \end{bmatrix}, \quad E = \begin{bmatrix} -6 \\ 1 \\ 2 \\ -4 \end{bmatrix}.$$

Then, construct the polyhedral convex cone $C_{n+1}(S)$ in Lemma 2.2, and illustrate Theorem 2.3.

SOLUTION. We first write

$$AX - Et = [A \mid -E]\begin{bmatrix} X \\ t \end{bmatrix} = \begin{bmatrix} -3 & -2 & 6 \\ -1 & 1 & -1 \\ 1 & -2 & -2 \\ 1 & -8 & 4 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ t \end{bmatrix} \le \boldsymbol{O}, \quad t \ge 0.$$

Then, from Example 2.5, we have

$$B = \{(4,1,1),(4,3,-1),(20,9,13),(4,9,5)\},$$

and

$$B^{\angle} = A^{*}.$$

Therefore,

$$C_{n+1}(S) = B^{\angle} \cap \{(r,s,t) \text{ in } E_3 \mid t \ge 0\}.$$

Clearly, this is a polyhedral convex cone, and hence there must exist some set B_0 such that

$$B_0^{\angle} = C_{n+1}(S).$$

Geometrically, B_0 contains vectors of B except for (4,3,-1), whose last component < 0. Its remaining vectors, which in fact can be formally computed, are (1,1,0), and (2,1,0), so that

$$B_0 = \{(4,1,1),(20,9,13),(4,9,5),(1,1,0),(2,1,0)\}.$$

To illustrate Theorem 2.3, we first note that

$$C_{n+1}(S) = \{u_1\begin{bmatrix}4\\1\\1\end{bmatrix}+u_2\begin{bmatrix}\frac{20}{13}\\\frac{9}{13}\\1\end{bmatrix}+u_3\begin{bmatrix}\frac{4}{5}\\\frac{9}{5}\\1\end{bmatrix}+v_1\begin{bmatrix}1\\1\\0\end{bmatrix}+v_2\begin{bmatrix}2\\1\\0\end{bmatrix} \mid U, V \geq \boldsymbol{O}\}$$

if and only if

$$\sum_{j=1}^{3} u_j = 1.$$

It follows that

$$S = \{u_1\begin{bmatrix}4\\1\end{bmatrix}+u_2\begin{bmatrix}\frac{20}{13}\\\frac{9}{13}\end{bmatrix}+u_3\begin{bmatrix}\frac{4}{5}\\\frac{9}{5}\end{bmatrix} \mid U \geq \boldsymbol{O}, \sum_{i=1}^{3} u_i = 1\} + \{v_1\begin{bmatrix}1\\1\end{bmatrix}+v_2\begin{bmatrix}2\\1\end{bmatrix} \mid V \geq \boldsymbol{O}\}.$$

Therefore, $$S = P^{\Delta} + Q^{\angle},$$

where

$$P = \begin{bmatrix}4 & 1\\ \frac{20}{13} & \frac{19}{13}\\ \frac{4}{5} & \frac{9}{5}\end{bmatrix}, \quad Q = \begin{bmatrix}1 & 1\\2 & 1\end{bmatrix}.$$

CHAPTER III

DUAL LINEAR INEQUALITIES

This chapter aims to derive some basic theorems which will be used to develop main theorems for the dual linear programs in the next chapter. The main references of this chapter are Tucker [12], Goldman [5], and Goldman and Tucker [6].

DEFINITION 3.1. Let A, C be (given) mxn marices, and U, X be (unknown) m×1, n×1 matrices. Then, each of the following pairs of systems of linear inequalities is called a pair of **Dual Systems**:

$$A^TU \geq \mathbf{O}, \text{ and } AX = \mathbf{O},\ X \geq \mathbf{O};$$

$$C^TU \geq \mathbf{O},\ U \geq \mathbf{O}, \text{ and } -CX \geq \mathbf{O},\ X \geq \mathbf{O}.$$

Each pair of dual systems will possess solutions which will be stated in Theorem 3.2, and Theorem 3.3. To derive these theorems we first need the following lemma.

LEMMA 3.1. The dual systems

$$A^TU \geq \mathbf{O}, \text{ and } AX = \mathbf{O},\ X \geq \mathbf{O} \tag{3.1}$$

possess solutions U and X such that

$$A_1^TU + x_1 > 0. \tag{3.2}$$

Proof. We proceed by induction on n, the number of columns in A. The initial case n = 1 is trivial: First, we can take $U = A_1$. Then, if $A_1 = \mathbf{0}$, take

$$x_1 = 1. \tag{3.3}$$

Otherwise, take

$$x_1 = 0. \tag{3.4}$$

It follows, in both cases, that (3.1) and (3.2) are satisfied.

Now, assume that the lemma is true for a matrix A of n columns, and proceed to prove it for a matrix

$$\bar{A} = [A\ A_{n+1}] = [A_1\ A_2\ ..\ A_{n+1}].$$

Applying the lemma to A, we get U_0 and X_0 that satisfy (3.1) and (3.2). If $A_{n+1}{}^T U_0 \geq 0$, take

$$U = U_0, \text{ and } X = [X_0{}^T\ 0]^T. \tag{3.5}$$

Then, clearly, U and X satisfy (3.1) and (3.2). This extends the lemma to the matrix $\bar{A}$.

If $A_{n+1}{}^T U_0 < 0$, apply the lemma to the matrix B, defined by

$$B = [B_1\ B_2\ ..\ B_n] = [A_1+\lambda_1 A_{n+1}\ A_2+\lambda_2 A_{n+1}\ ..\ A_n+\lambda_n A_{n+1}], \tag{3.6}$$

where

$$\lambda_j = -A_j{}^T U_0\ /\ A_{n+1}{}^T U_0 \geq 0,\ j = 1,2,..,n. \tag{3.7}$$

Note that

$$B_j^T U = (A_j - \frac{A_j^T U}{A_{n+1}^T U} A_{n+1})^T U = A_j^T U - A_j^T U = 0.$$

It follows that

$$B^T U = 0. \tag{3.8}$$

Using the lemma again, it yields V_0 and Y_0 such that

$$B^TV_0 \geq \mathbf{O},\ BY_0 = \mathbf{O},\ Y_0 \geq \mathbf{O},\ \text{and } B_1{}^TV_0 + y_1 > 0. \tag{3.9}$$

Take

$$Y = (Y_0, \textstyle\sum_{1\leq j\leq n}\lambda_j y_j). \tag{3.10}$$

Since

$$Y_0 \geq \mathbf{O},\ \text{and } \textstyle\sum_{1\leq j\leq n}\lambda_j y_j \geq 0,$$

we have

$$Y \geq \mathbf{O},$$

and

$$\begin{aligned}\bar{A}Y &= [A\ A_{n+1}]Y = \textstyle\sum_{1\leq j\leq n+1}A_jy_j = \sum_{1\leq j\leq n}A_jy_j + A_{n+1}\sum_{1\leq j\leq n}\lambda_j y_j \\ &= \textstyle\sum_{1\leq j\leq n}A_jy_j + \sum_{1\leq j\leq n}(\lambda_j A_{n+1})y_j = \sum_{1\leq j\leq n}(A_j + \lambda_j A_{n+1})y_j \\ &= [A_1+\lambda_1A_{n+1}\ \ A_2+\lambda_2A_{n+1}\ \ ..\ \ A_n+\lambda_nA_{n+1}]Y_0 = BY_0 = \mathbf{O}.\end{aligned}$$

Now, define

$$W = V + \mu U, \tag{3.11}$$

where

$$\mu = -(A_{n+1}{}^TV)/(A_{n+1}{}^TU). \tag{3.12}$$

Then,

$$A_{n+1}{}^TW = A_{n+1}{}^TV + \mu A_{n+1}{}^TU = A_{n+1}{}^TV - (A_{n+1}{}^TV/A_{n+1}{}^TU)A_{n+1}{}^TU = \mathbf{O}.$$

Since

$$B_j{}^TW = (A_j{}^T + \lambda_j A_{n+1}{}^T)W = A_j{}^TW + \lambda_j A_{n+1}{}^TW = A_j{}^TW,$$

we have

$$\bar{A}^TW = B^TW.$$

Also,

$$B^TU = \mathbf{O}.$$

We thus have

$$A^TW = B^TV + \mu B^TU = B^TV \geq \mathbf{O}.$$

With (3.9), this implies that

$$A_1^T W + y_1 = B_1^T V + y_1 > 0.$$

Hence, with W and Y from (3.11) and (3.10), we extend the lemma to the matrix $\bar{A} = [A\ A_{n+1}]$. The induction on n is fully established and the lemma holds for all n. ∎

EXAMPLE 3.1. Find the solutions U and X for (3.1) such that

$$A_1^T U + x_1 > 0 \text{ (likewise for } \bar{A}_1)$$

for each of the following cases:

$$a) \quad A = \begin{bmatrix} -3 \\ -1 \\ 1 \\ 1 \end{bmatrix}, \qquad b) \quad \bar{A} = \begin{bmatrix} -3 & -2 \\ -1 & 1 \\ 1 & -2 \\ 1 & -8 \end{bmatrix}.$$

SOLUTION. For (a), since $A \neq \mathbf{O}$, using (3.4), we can take

$$U = A \text{ and } X = \mathbf{O}.$$

For (b), we first set

$$\bar{A} = [A_1\ A_2] = [A\ A_2].$$

Using part (a) for A, we obtain

$$U = (-3, -1, 1, 1), \text{ and thus } A_2^T U = -5 < 0.$$

we next compute

$$\lambda_1 = 12/5 \text{ (by (3.7))},$$

and

$$B = (-39/5,\ 7/5,\ -19/5,\ -9/5) \text{ (by (3.6))}.$$

Now, using (3.4) for B, we have

$$V_0 = B, \text{ and } Y_0 = \mathbf{O}.$$

Hence, by (3.10), $Y = (0,0)$. By (3.12), and (3.11), we have

$$\mu = 851/25, \text{ and } W = 1/25(-2748,-816,756,396).$$

As a result,

$$U = W, \text{ and } X = Y$$

are desired solutions.

THEOREM 3.2. The dual systems

$$A^T U \geq \mathbf{O}, \text{ and } AX = \mathbf{O},\ X \geq \mathbf{O} \tag{3.17}$$

possess solutions U and X such that

$$A^T U + X > \mathbf{O}. \tag{3.18}$$

PROOF. For each $j = 1,2,..,n$, there exists corresponding solutions U^j, X^j such that

$$A^T U^j \geq \mathbf{O},\ AX^j = \mathbf{O},\ X^j \geq \mathbf{O},\ A_j^T U^j + x^j_j > 0.$$

By taking

$$U = \sum_{1\leq j\leq n} U^j, \text{ and } X = \sum_{1\leq j\leq n} X^j, \tag{3.19}$$

we have

$$A^T U = \sum_{1\leq j\leq n} A^T U^j \geq \mathbf{O}, \text{ and } AX = \sum_{1\leq j\leq n} AX^j = \mathbf{O}.$$

It follows that

$$A_j^T U + x_j = A_j^T (\sum_{1\leq k\leq n} U^k) + \sum_{1\leq k\leq n} X^k_j = \sum_{1\leq k\leq n} (A_j^T U^k + X^k_j),$$

and

$$A_j^T U^k + X^k_j \geq 0 \text{ (since } A^T U^k \geq \mathbf{O} \text{ and } X^k \geq \mathbf{O}.)$$

Hence, we obtain

$$A_j^T U + x_j \geq A_j^T U^j + x^j_j > 0.$$

This immediately implies that

$$A^T U + X > \mathbf{O}. \quad \blacksquare$$

The next two examples are given to illustrate the preceding theorem.

EXAMPLE 3.2. Let

$$A = \begin{bmatrix} -3 & -2 \\ -1 & 1 \\ 1 & -2 \\ 1 & -8 \end{bmatrix},$$

find solutions U and X for (3.17) that satisfy (3.18).

SOLUTION. From Example 3.1, we have

$U^1 = 1/25(-2748,-816,756,396)$, and $X^1 = (0,0)$.

To find U^2, X^2, let

$$A = \begin{bmatrix} -2 & -3 \\ 1 & -1 \\ -2 & 1 \\ -8 & 1 \end{bmatrix}.$$

Then, by using (3.4), we can take

$$U = [-2 \ 1 \ -2 \ 8]^T, \text{ and } X = \mathbf{O}.$$

Since

$$[-3 \ -1 \ 1 \ 1]U = -5 < 0,$$

we obtain, via (3.7),

$$\lambda_1 = 73/5, \text{ and } B = 1/5(229,-68,63,33).$$

Using (3.4) for B, we have

$$V = B, \text{ and } Y = \mathbf{O}.$$

By (3.12), (3.11),

$$\mu = 523/25, \text{ and } W = 1/25(99,183,-731,-4019).$$

Therefore,

$$U^2 = W, \text{ and } X^2 = Y.$$

By (3.19), we compute

$$U = 1/25(-2649,-633,25,-3623), \text{ and } X = \mathbf{O}.$$

COROLLARY 3.2.1. Let U^*, X^* be be solutions of of Theorem 3.2. Then, the (new) dual systems

$$[\mathbf{I}\ A]^T U \geq \mathbf{O}, \text{ and } [\mathbf{I}\ A]X = \mathbf{O},\ X \geq \mathbf{O},$$

possess solutions U^* and $[\mathbf{O}\ X^{*T}]^T$ such that

$$[\mathbf{I}\ A]^T U^* + X^* > \mathbf{O}, \text{ provided that } U^* \geq \mathbf{O}.$$

PROOF. From Theorem 3.2, we have

$$[\mathbf{I}\ A]^T U^* \geq \mathbf{O},$$

$$[\mathbf{I}\ A]\,[\mathbf{O}\ X^{*T}]^T = \mathbf{IO} + AX^* - AX^* = \mathbf{O},$$

and

$$[\mathbf{O}\ X^{*T}]^T \geq \mathbf{O}. \quad \blacksquare$$

EXAMPLE 3.3. To illustrate corollary 3.2.1, let

$$a)\ A = \begin{bmatrix} 2 \\ 2 \\ 1 \\ 1 \end{bmatrix}, \quad b)\ \bar{A} = \begin{bmatrix} 2 & 1 \\ 2 & 0 \\ 1 & 1 \\ 1 & 2 \end{bmatrix}, \quad c)\ E = \begin{bmatrix} 1 \\ 0 \\ 1 \\ 2 \end{bmatrix}, \quad d)\ \bar{E} = \begin{bmatrix} 1 & 2 \\ 0 & 2 \\ 1 & 1 \\ 2 & 1 \end{bmatrix}, \quad e)\ C = \bar{A}.$$

In each case, find solutions U^* and X^* such that

$A_1^T U^* + x^*_1 > 0$ (likewise for $\bar{A}_1$, E_1, $\bar{E}_1$) $(C^T U^* + X^* > \mathbf{O})$.

SOLUTION. a) Since $A \neq \mathbf{O}$, we can take

$$U^* = A, \text{ and } X^* = \mathbf{O}.$$

b) Let $\bar{A} = [A_1\ A_2] = [A\ A_2]$. From part (a),

$$U = A, \text{ and hence } A_2^T U = 5 \geq 0.$$

It follows from (3.5) that

$$X^* = [X^T\ 0]^T = \mathbf{O}, \text{ and } U^* = U.$$

c) Since $E \neq \mathbf{O}$, we can take

$$U^* = E, \text{ and } X^* = \mathbf{O}.$$

d) Let $\bar{E} = [E_1\ E_2] = [E\ E_2]$. From part (c),

$$U = E, \text{ and hence } E_2^T U = 5 \geq 0.$$

From (3.5),

$$X^* = [X^T\ 0]^T = \mathbf{O}, \text{ and } U^* = U.$$

e) From (b) and (c), we obtain

$$U^1 = A,\ X^1 = \mathbf{O},\ U^2 = E,\ X^2 = \mathbf{O}.$$

It follows from (3.19) that

$$U^* = U^1 + U^2 = [3\ 2\ 2\ 3]^T, \text{ and } X^* = X^1 + X^2 = \mathbf{O}.$$

THEOREM 3.3. The dual systems

$$U \geq \mathbf{O},\ C^T U \geq \mathbf{O} \text{ and } -CY \geq \mathbf{O},\ Y \geq \mathbf{O} \tag{3.20}$$

possess solutions U^*, and Y^* such that

$$U^* - CY^* > \mathbf{O} \text{ and } C^T U^* + Y^* > \mathbf{O}. \tag{3.21}$$

PROOF. Let $A = [\mathbf{I}\ C]$. Then, by Theorem 3.2, there exist solutions U^*, and X^* for (3.17) such that $A^T U^* + X^* > \mathbf{O}$. Write

$$X = [W^T \; Y^T]^T, \text{ and } X^* = [W^{*T} \; Y^{*T}]^T.$$

Then,

$$[\mathbf{I} \; C]^T U^* \geq 0, \quad [\mathbf{I} \; C][W^{*T} \; Y^{*T}]^T = \mathbf{O}, \quad [W^{*T} \; Y^{*T}]^T \geq \mathbf{O}. \tag{3.22}$$

Moreover,

$$[\mathbf{I} \; C]^T U^* + [W^{*T} \; Y^{*T}]^T > \mathbf{O},$$

and thus

$$U^* + W^* > \mathbf{O}, \text{ and } C^T U^* + Y^* > \mathbf{O}. \tag{3.23}$$

It follows that

$$\mathbf{O} \leq \mathbf{I}^T U^* = \mathbf{I} U^* = U^*, \text{ and } C^T U^* \geq \mathbf{O}. \tag{3.24}$$

Note that, from (3.22),

$$W^* = -CY^*. \tag{3.25}$$

Therefore, from (3.24), (3.25), (3.22), (3.23), we have

$$U^* \geq \mathbf{O}, \; C^T U^* \geq \mathbf{O} \text{ and } -CY^* = W^* \geq \mathbf{O}, \; Y^* \geq \mathbf{O},$$

such that

$$U^* - CY^* > \mathbf{O}, \; C^T U^* + Y^* > \mathbf{O}.$$

EXAMPLE 3.4. Find solutions U^*, and Y^* for (3.20), and (3.21), where

$$C = \begin{bmatrix} 2 & 1 \\ 2 & 0 \\ 1 & 1 \\ 1 & 2 \end{bmatrix}.$$

SOLUTION. Write $A = [\mathbf{I} \; C]$. Then, from Example 3.3(e), and corollary 3.2.1, there exist solutions

$$U^* = [3 \; 2 \; 2 \; 3]^T, \text{ and } X^* = \mathbf{O}$$

for (3.17) such that

$$A^TU^* + X^* > \mathbf{O}.$$

So, we can find

$$W^* = \mathbf{O}, \text{ and } Y^* = \mathbf{O}$$

such that

$$[\mathbf{I}\ C]^TU^* + [W^{*T}\ Y^{*T}]^T > \mathbf{O},$$

and thus

$$U^* + W^* > \mathbf{O}, \text{ and } CU^* + Y^* > \mathbf{O}. \tag{3.26}$$

From (3.17),

$$[\mathbf{I}\ C]\,[W^{*T}\ Y^{*T}]^T = \mathbf{O}, \text{ so that } W^* = -CY^*.$$

Hence, by (3.26),

$$U^* - CY^* > \mathbf{O} \text{ and } CU^* + Y^* > \mathbf{O}.$$

In computation, we have

$$\begin{bmatrix}3\\2\\2\\3\end{bmatrix} - \begin{bmatrix}2&1\\2&0\\1&1\\1&2\end{bmatrix}\begin{bmatrix}0\\0\end{bmatrix} = \begin{bmatrix}3\\2\\2\\3\end{bmatrix} > \mathbf{O},$$

and

$$\begin{bmatrix}2&2&1&1\\1&0&1&2\end{bmatrix}\begin{bmatrix}3\\2\\2\\3\end{bmatrix} + \begin{bmatrix}0\\0\end{bmatrix} = \begin{bmatrix}15\\11\end{bmatrix} > \mathbf{O}.$$

DEFINITION 3.2. Let K be a skew matrix, i.e., $K^T = -K$. Then, the following system

$$KW \geq \mathbf{O}, \text{ and } W \geq \mathbf{O} \tag{3.27}$$

is called a **Self Dual System**.

THEOREM 3.4. The self dual system, defined above, has a solution W^* such that

$$KW^* + W^* > \mathbf{O}. \tag{3.28}$$

PROOF. Let $C = K^T$. Then,

$$C^T = K = -K^T = -C.$$

And hence, we have the dual system

$$C^TU \geq \mathbf{O},\ U \geq \mathbf{O},\ \text{and}\ -CY \geq \mathbf{O},\ Y \geq \mathbf{O}. \tag{3.29}$$

By Theorem 3.3, there exist solutions U^*, and Y^* such that

$$U^* - CY^* > \mathbf{O}\ \text{and}\ C^TU^* + Y^* > \mathbf{O}. \tag{3.30}$$

It follows that

$$U^* \geq \mathbf{O},\ KU^* \geq \mathbf{O},\ KY^* \geq \mathbf{O},\ Y^* \geq \mathbf{O},$$

$$U^* + KY^* > \mathbf{O},\ \text{and}\ C^TU^* + Y^* > \mathbf{O}.$$

Therefore,

$$K(U^* + Y^*) \geq \mathbf{O},\ U^* + Y^* \geq \mathbf{O},\ K(U^* + Y^*) + (U^* + Y^*) \geq \mathbf{O}.$$

It follows that

$$W^* = U^* + Y^*$$

is a desired solution. ■

EXAMPLE 3.5. Let

$$K = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix},$$

find a solution W^* that satisfies (3.28).

SOLUTION. First, we need to find solutions U^*, and Y^* for (3.29), and (3.30), where

$$C = [C_1\ C_2] = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}.$$

Since $C_1 = [0\ 1]^T$, we take

$$U_1 = C_1,\ X_1 = \mathbf{O}.$$

Then,

$$C_2^T U_2 = 0.$$

It follows from (3.5) that

$$U^2 = U_2,\ X^2 = [X_2^T\ 0]^T = \mathbf{O}.$$

Therefore,

$$U^* = U^1 + U^2 = [-1\ 1]^T,\ \text{and}\ X^* = [0\ 0]^T$$

are desired solutions for (3.29), and (3.30). It follows that

$$W^* = U^* + X^* = [-1\ 1]^T$$

is a solution for (3.27), and (3.28).

The following corollary will be used in chapter IV to prove important theorems.

COROLLARY 3.4.1. Let

$$K = -\begin{bmatrix} \mathbf{O} & A & -B \\ -A^T & \mathbf{O} & C \\ B^T & -C^T & \mathbf{O} \end{bmatrix},$$

where A, B, C are mxn, mx1, nx1 matrices, respectively, and

$\mathbf{O}$'s are square matrices whose sizes agree with other matrices in the block matrix. Then, K is clearly a skew symmetric matrix, and the self dual system (3.27) will have a solution $[U_o^T\ X_o^T\ t_o]^T$ such that

$$Bt_0 \geq AX_0, \tag{3.31}$$

$$A^TU_0 \geq Ct_0, \tag{3.32}$$

$$C^TX_0 \geq B^TU_0, \tag{3.33}$$

$$C^TX_0 + t_0 > B^TU_0, \tag{3.34}$$

$$U_o + Bt_0 > AX_o, \tag{3.35}$$

$$A^TU_0 + X_o > Ct_0, \tag{3.36}$$

where U_0, and X_0 are mx1, nx1 matrices, respectively, and t_0 is a scalar.

PROOF. By applying Theorem 3.4, the self dual system posseses a solution W^* such that $KW^* + W^* > \mathbf{O}$. Write

$$[U_o^T\ X_o^T\ t_o] = W^{*T}.$$

Then,

$$-\begin{bmatrix} \boldsymbol{O}U_0+AX_0-Bt_0 \\ -A^TU_0+\boldsymbol{O}X_0+Ct_0 \\ B^TU_0-C^TX_0+\boldsymbol{O}t_0 \end{bmatrix} = KW^* \geq \boldsymbol{O}.$$

This proves (3.31)-(3.33). Moreover,

$$-\begin{bmatrix} \boldsymbol{O}U_0+AX_0-Bt_0 \\ -A^TU_0+\boldsymbol{O}X_0+Ct_0 \\ B^TU_0-C^TX_0+\boldsymbol{O}t_0 \end{bmatrix} + \begin{bmatrix} U_0 \\ X_0 \\ t_0 \end{bmatrix} > \boldsymbol{O}$$

is equivalent to

$$-\begin{bmatrix} B^{\top}U_0 - C^{\top}X_0 \\ AX_0 - Bt_0 \\ -A^{\top}U_0 + Ct_0 \end{bmatrix} + \begin{bmatrix} t_0 \\ U_0 \\ X_0 \end{bmatrix} > \mathbf{O},$$

which proves (3.34)-(3.36). ■

EXAMPLE 3.6. In each of the following matrices:

$$a)\ A = \begin{bmatrix} 0 & 1 & -1 \\ -1 & 0 & -1 \\ 1 & 1 & 0 \end{bmatrix},\ b)\ A = \begin{bmatrix} 1 & 0 & -1 \\ 0 & -1 & -1 \\ 1 & 1 & 0 \end{bmatrix},\ c)\ A = \begin{bmatrix} -1 & 1 & 0 \\ -1 & 0 & -1 \\ 0 & 1 & 1 \end{bmatrix},$$

find solutions U, and X such that (3.2) holds true. Use these results to compute U^0, X^0 for (3.17) and (3.18), where A is defined as in (a).

SOLUTION. a) Let $A = [A_1\ A_2\ A_3]$. Since $A_1 \neq \mathbf{O}$, we can take

$$U_1 = A_1, \text{ and } X_1 = \mathbf{O}.$$

Then,

$$A_2^{\top}U_1 = 1 \geq 0.$$

So by (3.7),

$$U_2 = U_1 \text{ and } X_2 = [X_1^{\top}\ 0]^{\top}.$$

Next, compute

$$A_3^{\top}U_2 = 1 \geq 0.$$

So, we can take, by (3.7),

$$U_3 = U_2, \text{ and } X_3 = [X_2^{\top}\ 0]^{\top}.$$

Hence,

$$U\ (= U_3) = A_1, \text{ and } X\ (= X_3) = \mathbf{O}.$$

b) Since $A_1 \neq \mathbf{O}$, we take

$$U_1 = A_1, \text{ and } X_1 = \mathbf{O}.$$

Then,

$$A_2^T U_1 = 1 \geq 0.$$

So, by (3.5), take

$$U_2 = U_1 \text{ and } X_2 = [X_1^T\ 0]^T.$$

Since

$$A_3^T U_2 = -1 < 0,$$

by (3.9), (3.8),

$$\lambda_1 = -A_1^T U_2 / -A_3^T U_2 = 2 \geq 0,$$

$$\lambda_2 = -A_2^T U_2 / -A_3^T U_2 = 1 \geq 0,$$

$$B_1 = A_1 + \lambda_1 A_3 = [-1\ -2\ 1]^T,$$

$$B_2 = A_2 + \lambda_2 A_3 = [-1\ -2\ 1]^T.$$

Now, repeat the process for $B = [B_1\ B_2]$. Since $B_1 \neq \mathbf{O}$, we can take

$$V_1 = B_1, \text{ and } Y_1 = O.$$

Since

$$B_2^T V_1 = 6 \geq 0,$$

by (3.5), we obtain

$$V_2 = V_1, \text{ and } Y_2 = [Y_1^T\ 0]^T.$$

By (3.10), we compute

$$\mu = -A_3^T V / -A_3^T U_2 = 3,$$

$$U = A_1 + \lambda_1 A_3 = [-1\ -2\ 1]^T,$$

$$X = [Y^T\ 0]^T = \mathbf{O}.$$

c) Similar to previous parts, we first obtain

$$\lambda_1 = -A_1^T U_1 / -A_2^T U_1 = 2 \geq 0,$$

$$B_1 = A_1 + \lambda_1 A_2 = [1\ -1\ 2]^T.$$

So, by repeating the process for B_1, we compute:

$$\mu = -A_2^T V / -A_2^T U_1 = 3,$$

$$U_2 = V + \mu U_1 = [-2\ -4\ 2]^T,$$

$$X_2 = [Y^T\ 0]^T = \mathbf{0}.$$

$$A_3^T U_2 = 6 \geq 0.$$

Therefore, by (3.5), we can have

$$U = U_2, \text{ and } X = [X_2^T\ 0]^T = \mathbf{0}.$$

Finally, from (a)-(c), via (3.19),

$$U^0 = [0\ -7\ 7]^T, \text{ and } X^0 = \mathbf{0}.$$

EXAMPLE 3.7. (Related to Corollary 3.4.1) Let $A = [1]$, $B = [1]$, and $C = [-1]$. Then, compute $[U_o^T\ X_o^T\ t_o]^T$ that satisfies (3.31)-(3.36).

SOLUTION. Let

$$K = -\begin{bmatrix} 0 & 1 & -1 \\ -1 & 0 & -1 \\ 1 & 1 & 0 \end{bmatrix}.$$

Then, by Example 3.6, we obtain

$$U^0 = [0\ -7\ 7]^T, \text{ and } X^0 = \mathbf{0}$$

that satisfy (3.17), (3.18) with $A = -K$. Following the proof of Theorem 3.3, with some computation, we find

$$U^* = [0\ 0\ 7]^T, \text{ and } X^* = [1\ 1\ 0\ 0\ 0\ 1\ ]^T$$

that satisfy (3.18) with $A = [\mathbf{I}\ -K]$. From this, we have

$$U^* = [0\ 0\ 7]^T, \text{ and } Y^* = [0\ 0\ 1\ ]^T,$$

which satisfy (3.20) and (3.21). It follows from the proof of Theorem 3.4 that

$$W^* = U^* + Y^* = [0\ 0\ 8\ ]^T$$

is a desired solution that satisfies (3.28). Therefore, from Corollary 3.4.1, we can take

$$U_0 = [0],\ X_0 = [0],\ t_0 = 8,$$

that satisfy (3.31)-(3.36):

$$Bt_0 = 8 \geq AX_0 = 0,$$

$$A^TU_0 = 0 \geq Ct_0 = -8,$$

$$C^TX_0 = 0 \geq B^TU_0 = 0,$$

$$C^TX_0 + t_0 = 8 > B^TU_0 = 0,$$

$$U_0 + Bt_0 = 8 > AX_0 = 0,$$

$$A^TU_0 + X_0 = 0 > Ct_0 = -8.$$

Also, it can be checked that

$$KW \geq \mathbf{O},\ W \geq \mathbf{O}.$$

LEMMA 3.2. Let A, A_0 be mxn, mx1 matrices, respectively. suppose that for each U with $A^TU \geq \mathbf{O}$, $A_0^TU \geq 0$. Then, there exists an n-vector $X_0 \geq \mathbf{O}$ such that $A_0 = AX_0$.

PROOF. Consider the dual systems

$$[-A_0\ A]^TU \geq \mathbf{O} \text{ and } [-A_0\ A][x_0\ X^T]^T = \mathbf{O},\ [x_0\ X^T]^T \geq \mathbf{O}. \quad (3.37)$$

Then, by Lemma 3.1, there exist vectors U and $[x_0\ X^T]^T$ such that

$$-A_0^TU + x_0 > 0. \tag{3.38}$$

Furthermore, from (3.37), we also have

$$-A_0^TU \geq 0,\ A^TU \geq \mathbf{O},\ -A_0x_0 + AX = \mathbf{O},\ x_0 \geq 0,\ X \geq \mathbf{O}. \tag{3.39}$$

Therefore, by the hypothesis,

$$-A_0^TU \geq 0 \text{ (since } A^TU \geq \mathbf{O}).$$

It follows from (3.39) that

$$x_0 > A_0^TU \geq 0.$$

Let

$$X_0 = x_0^{-1}X.$$

Then, by (3.39),

$$-A_0x_0 + A(x_0X_0) = \mathbf{O},$$

i.e.,

$$A_0 = AX_0. \quad \blacksquare$$

COROLLARY 3.2.1. Let A, Y be mxn and nx1 matrices, respectively. Suppose that for each vector X with $AX \leq \mathbf{O}$, $Y^TX \leq 0$. Then,

$$Y = \textstyle\sum_{1\leq i\leq m} u_iA_i,$$

for some u_i's ≥ 0. Here, A_i's denote row vectors of A.

PROOF. Let

$$B = -A^T, \text{ and } B_0 = -Y.$$

Then, B is an nxm matrix and $-B^T = A$. Therefore, by hypothesis, $B^TX \geq \mathbf{O}$ implies that $B_0^TX \geq 0$. By Lemma 3.2, there exists an m-vector $U = [u_1\ u_2\ ..\ u_m]^T$ such that $B_0 = BU$. It follows that

$$-Y = (-A^T)U,$$

i.e.,

$$Y = A^TU = \sum_{1\le i\le m} u_iA_i. \quad ■$$

We next state and prove a theorem that characterizes polar of a polyhedral convex cone in terms of convex cone hull.

THEOREM 3.4. (**Farcas' Theorem**) For each mxn matrix A,

$$A^{**} = (A^T)^{\angle}.$$

PROOF. Let $Y \in A^{**}$. Then, by Definition 3.3,

$$Y^TX \le 0 \text{ for each } X \in A^*.$$

So, by Lemma 3.2, we have

$$Y = \sum_{1\le i\le m} u_iA_i \text{ for } u_i\text{'s} \ge 0.$$

It follows from (2.12) that

$$Y \in (A^T)^{\angle}.$$

For the reverse inclusion, let $Y \in A^{\angle}$. Suppose

$$Y = \sum_{1\le i\le m} u_iA_i \text{ for } u_i\text{'s} \ge 0.$$

Then, for each $X \in A^{**}$,

$$AX \le \mathbf{0},$$

i.e.,

$$A_i^TX \le 0 \text{ for } i = 1,2,..,m.$$

Since

$$Y^TX = (\sum_{1\le i\le m} u_iA_i^T)X = \sum_{1\le i\le m} u_i(A_i^TX) \le 0,$$

it follows that

$$Y \in A^{**}. \quad \blacksquare$$

Theorem 3.4 implies that

$$(A^*)^\angle = ((A^T)^\angle)^*,$$

noting that

$$(A^*)^\angle = A^*,$$

and

$$((A^T)^\angle)^* = (A^{**})^* = A^*.$$

THEOREM 3.5. (**Weyl's Theorem**) For each mxn matrix, there exists a finite set B of vectors in E_n such that

$$A^* = B^\angle, \text{ and } (A^T)^\angle = (B^T)^*.$$

PROOF. That $A^* = B^\angle$ was already covered by Theorem 2.2. To prove that $(A^T)^\angle = (B^T)^*$, we first observe that

$$(B^\angle)^* = (B^T)^*.$$

So, By the first part,

$$A^{**} = (B^T)^*.$$

It follows from Theorem 3.4 that

$$(A^T)^\angle = (B^T)^*. \quad \blacksquare$$

Theorem 3.5 immediately implies the following corollary, that can be considered as a geometric version of the theorem.

COROLLARY 3.5.1. The convex hull of finitely many half lines is the intersection of finitely many half spaces, and conversely.

LEMMA 3.3. For each vector X_0 and each polyhedral convex cone $(A^T)^\angle = \{A^TV \mid V \geq \mathbf{0}\}$, either $X \in (A^T)^\angle$ or there exists a hyperplane $Y^TX = 0$ that separates X_0 from $(A^T)^\angle$, i.e., X_0 lies in the open half space $X^TY > 0$ and $(A^T)^\angle$ lies in the closed half space $X^TY \leq 0$.

PROOF. By Theorem 3.4, we have

$$(A^T)^\angle = A^{**} = \{X \in E_n \mid X^TY \leq 0 \text{ for all } Y \in A^*\}.$$

Suppose $X \notin (A^T)^\angle$. Then, $X_0 \notin A^{**}$, and hence there exists some vector Y in A^* such that $Y^TX_0 > 0$. On the other hand, we have

$$Y^TX \leq 0 \text{ for all } X \text{ in } (A^T)^\angle.$$

It follows that the hyperplane Y^TX separates X_0 and $(A^T)^\angle$. ■

THEOREM 3.6. (**Separation Theorem**) Let

$$P^\vartriangle = \{PU \mid U \geq \mathbf{0},\ \textstyle\sum_{1\leq i\leq m} u_i = 1\},$$

and

$$Q^\angle = \{QV \mid V \geq \mathbf{0}\}$$

be a bounded convex polyhedron and a polyhedral convex cone, respectively. Then, either $P^\vartriangle \cap Q^\angle \neq \emptyset$, or there exists a hyperplane $X^TY = 0$ that separates $P^\vartriangle$ and $Q^\angle$, i.e., $P^\vartriangle$ lies in the open half space $X^TY > 0$, and $Q^\angle$ lies in the closed half space $X^TY \leq 0$.

PROOF. Let $P = \{P_1, P_2, .., P_p\}$. If $p = 1$, then, by Lemma 3.3, the desired results will follow. So, assume that $p > 1$, and suppose that $P^{\Delta} \cap Q^{\angle} = \emptyset$.

We first prove that for each $i = 1,2,..,p$,

$$p_i \notin S_i = \{-P_1, -P_2, .., -P_{i-1}, -P_{i+1}, .., P_p, Q_1, Q_2, .., Q_q\}.$$

Suppose the contrary that

$$P_i = \sum_{k \neq i} u_k(-P_k) + \sum_{1 \leq j \leq n} v_j Q_j, \text{ for some } u_k\text{'s, } v_j\text{'s} \geq 0.$$

Then,

$$\sum_{k \neq i} u_k P_k + P_i = \sum_{1 \leq j \leq n} s_j Q_j \text{ for some } s_j\text{'s}.$$

Let

$$r_k = u_k/a \text{ for } k \neq i, \ r_i = 1/a,$$
$$s_j = v_j/a \text{ for each } j,$$

where

$$a = 1 + \sum_{i \leq k \leq m} u_k.$$

We then have

$$\sum_k r_k P_k = \sum_{1 \leq j \leq n} s_j Q_j,$$

for some

$$u_k\text{'s} \geq 0, \ \sum_{i \leq k \leq m} u_k = 1, \ v_j\text{'s} \geq 0.$$

But this is impossible since P^{Δ} and $Q^{\angle}$ are disjoint. Therefore,

$$P_i \notin S_i.$$

It follows from Lemma 3.3 that, for each $i = 1,2,..,p$, there exists a vector Y_i such that

$$Y_i^T P_i > 0 \text{ and } Y_i^T X \leq 0 \text{ for each } X \in S_i.$$

This implies that for each $k = 1,2,..,p$, and each $X \in Q^{\angle}$,

$$Y_i^T P_i > 0 \text{ and } Y_i^T P_k \geq 0 \text{ for } k \neq i, \text{ and } Y_i^T X \leq 0.$$

Let

$$Y = \sum_{1 \le i \le p} Y_i .$$

Then, for each $i = 1,2,..,p$,

$$Y^T P_i = Y_i^T P_i + \sum_{k \ne i} Y_k^T P_k > 0 .$$

Moreover, it is clear that

$$Y^T X = \sum_{1 \le i \le p} Y_i^T X \text{ for each } X \in Q^{\not\,} . \quad \blacksquare$$

CHAPTER IV

LINEAR PROGRAMS

The object of this chapter is to present a linear programming problem together with its basic properties, and the simplex method that has been used to solve a linear program. In particular, extreme points and extreme directions will be treated as main concepts from which the simplex algorithm will be derived. The main references of the chapter are Bazaraa, Sherali and Shetty [1], Cooper and Steinberg [2], and Noble and Daniel [10].

DEFINITION 4.1. (**Linear Program in Inequality Form**) Let A and C be mxn and nx1 matrices, respectively. If there exists a vector $X_* \in E_n$ such that

$$\text{Inf } C^TX = C^TX_*,\ AX_* \le B, \text{ and } X_* \ge \mathbf{O},$$

then the problem

$$\text{Minimize} \qquad C^TX \tag{4.1}$$

$$\text{subject to} \qquad AX \le B, \tag{4.2}$$

$$\text{and} \qquad X \ge \mathbf{O}, \tag{4.3}$$

will be called an **Optimal Linear Program**. Otherwise, it will be called a **Nonoptimal Linear Program**. A **Linear Program** will be meant for either of these programs.

Replacing Minimize by Maximize, or $\le$ by $\ge$ (or $=$) in any inequality of (4.1), or even dropping (4.3), we obtain another kind of linear program which is, indeed, reducible to the Standard Form (4.1)-(4.3) above.

Moreover, replacing C, X by

$$\begin{bmatrix} C \\ \mathbf{O} \end{bmatrix}, \quad \begin{bmatrix} X \\ X_s \end{bmatrix},$$

respectively, and (4.2) by

$$AX + \mathbf{I}X_s = B, \; X_s \geq \mathbf{O},$$

where X_s is an nx1 matrix, we have a new standard form, called a **Slack Variable Form**, for a linear program.

Now, since

$$AX + \mathbf{I}X_s = [A \;\; \mathbf{I}] \begin{bmatrix} X \\ X_s \end{bmatrix},$$

we are prompted to define a linear program in terms of linear equality constraints.

DEFINITION 4.2. (**Linear Program in Equality Form**) Let A and C be defined as in Definition 4.1. If there exists a vector $X_* \in E_n$ such that

$$\text{Inf (or Sup) } C^TX = C^TX_*, \; AX_* = B, \; X_* \geq \mathbf{O},$$

then the problem

$$\text{Minimize (or Maximize)} \quad C^TX \tag{4.4}$$

$$\text{subject to} \quad AX = B, \tag{4.5}$$

$$\text{and} \quad X \geq \mathbf{O}, \tag{4.6}$$

will be called a **Minimal** (or **Maximal**) **Linear Program.**

Otherwise, it will be called a **Nonoptimal Linear Program**. **Optimal** will be used to mean minimal or maximal, and a **Linear Program** will be used for an optimal or nonoptimal one.

As a matter of convenience, "Linear" may be dropped and "Program" may be replaced by "Problem". Minimal (or Maximal) is sometimes used for a general linear problem to be minimized (or maximized) regardless of its optimality. Moreover, by symmetry, a program may be represented by either a minimal or a maximal one, and will be assumed to be minimal for this chapter when the context is clear.

DEFINITION 4.3. (**Constraint Set**) (4.1) (or (4.4)), (4.2) (or (4.5)), and (4.3) (or (4.6)) will be respectively called the **Objective Function**, the **Constraints**, and the **Nonnegative Condition**. The set of points in E_n that satisfy the contraints and the nonnegative condition will be called the **Constraint Set**.

DEFINITION 4.4. (**Basic and Optimal Feasible Solution**) A vector $X_0 \geq \mathbf{O}$ in E_n is called a **Feasible Solution** if it satisfies the constraints of a linear program. A feasible solution will be respectively called **Basic**, or **Optimal** if it has at most m positive entries, or it is a vector X_* defined by Definition 4.1 or 4.2.

EXAMPLE 4.1. Consider the problem

$$\begin{array}{lrcl} \text{Minimize} & x_1 - 4x_2 & & \\ \text{Subject to} & x_1 - 2x_2 & \geq & -4, \\ & x_1 + x_2 & \leq & 3, \\ & x_1,\ x_2 & \geq & 0. \end{array}$$

Then, illustrate the main concepts given in Definitions 4.1 and 4.2. Guess the optimal point, if any, graphically.

SOLUTION. By rewriting $x_1 - 2x_2 \geq -4$ as $-x_1 + 2x_2 \leq 4$, we put the problem in the standard form (4.1)-(4.3):

$$\begin{array}{lrcl} \text{Minimize} & & C^T X & \\ \text{subject to} & AX & \leq & B, \\ & X & \geq & \mathbf{O}, \end{array}$$

where

$$C = \begin{bmatrix} 1 \\ -4 \end{bmatrix}, \quad A = \begin{bmatrix} -1 & 2 \\ 1 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 4 \\ 3 \end{bmatrix}, \quad X = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.$$

By adding slack variables, we put the problem in the standard form (4.4)-(4.6):

$$\begin{array}{lrcl} \text{Minimize} & & C^T X & \\ \text{subject to} & AX & = & B, \\ & X & \geq & \mathbf{O}, \end{array}$$

where

Graphically, the level curves of the objective function are the straight lines, whose values decrease along the vector $-C = [-1\ 4]^T$. Therefore, a reasonable guess for the optimal solution would be the point $[2/3\ \ 7/3]^T$, which is the

$$C = \begin{bmatrix} 1 \\ -4 \\ 0 \\ 0 \end{bmatrix}, \quad A = \begin{bmatrix} -1 & 2 & 1 & 0 \\ 1 & 1 & 0 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 4 \\ 3 \end{bmatrix}, \quad X = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}.$$

intersection of the lines defined by

$$x_1 - 2x_2 = -4,$$

and

$$x_1 + x_2 = 3.$$

The following lemma will be used to prove Lemma 4.2 that gives a connection between (4.2)-(4.3) and (4.5)-(4.6). we omit the proof since it is simple.

LEMMA 4.1. The constraint set of a linear program defined by (4.4)-(4.6) is a convex set.

DEFINITION 4.5. (**Extreme Point**) A vector $X \in S \subseteq E_n$ is called an **Extreme Point** (**Vector**) of S if it can not be written as a convex combination of any two distinct points in S.

Clearly, X is an extreme point of S if and only if it is not a mean of any two other vectors of S, i.e., if

$$X = 1/2\ (X_1 + X_2),\ \text{where } X_1,\ X_2 \in S,$$

then

$$X = X_1 \text{ or } X = X_2 \text{ (and hence } X = X_1 = X_2.)$$

LEMMA 4.2. For each nonvacuous polyhedral convex set

$$P = \{X \geq \mathbf{O} \mid AX \leq B\}, \tag{4.7}$$

the set

$$Q = \{[X^T\ S^T]^T \geq \mathbf{O} \mid AX + S = B\} \tag{4.8}$$

is a nonvacuous polyhedral convex set that has its extreme points corresponding to the extreme points of P.

PROOF. First, by writing

$$AX + S = [A\ \ \boldsymbol{I}] \begin{bmatrix} X \\ S \end{bmatrix},$$

we see that Q is the intersection of two polyhedral convex sets, so that it is also a polyhedral convex set.

Furthermore, it is clear that Q is nonvacuous since P is nonvacuous. Let X be an extreme point of P. Then, by (4.7), we have

$$AX \leq B.$$

Hence, there exists a vector $S \in E_m$ such that

$$AX + S = B.$$

This implies that $[X^T\ S^T]^T \in Q$.

To show that this point is also an extreme point of Q, let us suppose that

$$[X^T\ S^T]^T = 1/2\ ([X_1^T\ S_1^T]^T + [X_2^T\ S_2^T]^T), \tag{4.9}$$

where

$$[X_1^T\ S_1^T]^T,\ [X_2^T\ S_2^T]^T \in Q.$$

Then, we have that X_1, $X_2 \in P$, and

$$[X^T\ S^T] = [1/2[X_1^T\ X_2^T]\quad 1/2[S_1^T\ S_2^T]].$$

It follows that

$$X = 1/2(X_1 + X_2), \text{ and } S = 1/2(S_1 + S_2).$$

Since X is an extreme point of P, we must have

$$X = X_1 = X_2, \text{ and hence } S = S_1 = S_2.$$

By (4.9), we have

$$[X^T \ S^T] = [X_1^T \ S_1^T] = [X_2^T \ S_2^T].$$

Thus,

$$[X^T \ S^T]^T \text{ is an extreme point of } Q,$$

so that

$$[X^T/2 \ S^T/2]^T \in Q.$$

Next, let $[X^T \ S^T]^T$ be an extreme point of Q. Then, $X \in P$. Suppose that

$$X = 1/2(X_1 + X_2), \text{ where } X_1, X_2 \in P.$$

Then, there exist $S_1, S_2 \in E_m$ such that

$$[X_1^T \ S_1^T]^T, \ [X_2^T \ S_2^T]^T \in Q,$$

and hence

$$[X^T \quad (S_1 + S_2)^T]^T \in Q.$$

This implies that

$$S = S_1 + S_2.$$

By Lemma 4.1, Q is a convex set. It follows that

$$1/2([X_1^T \ S_1^T]^T + [X_2^T \ S_2^T]^T) = [1/2(X_1+X_2)^T \ 1/2(S_1 + S_2)^T]^T$$
$$= [X^T \ S^T]^T \in Q.$$

But $[X^T \ S^T]^T$ is an extreme point of Q. Therefore,

$$[X^T \ S^T] = [X_1^T \ S_1^T] = [X_2^T \ S_2^T].$$

We immediately have

$$X = X_1 = X_2,$$

and hence, X is an extreme point of P. ■

The next three theorems are fundamental and useful in proving many characterization theorems. Their proofs can be found in Cooper and Steinberg [2], or Bazaraa, Sherali and Shetty [1]. Also, when not specified, a linear program and its constraint set will be used in the sense of (4.4)-(4.6). we also note that a matrix A can be modified into a full rank one, so that without loss of generality we may assume that A is of rank m.

THEOREM 4.1. Let

$$S = \{X \in E_n \mid AX = B \text{ and } X \geq \mathbf{0}\} \neq \varnothing.$$

Suppose that $A_1, A_2, .., A_k$, $k \leq m$, are linear independent column vectors of A such that

$$x_1A_1 + x_2A_2 + .. + x_kA_k = B,$$

where all x_j's ≥ 0. Then, the vector

$$X_0 = (x_1, x_2, .., x_k, 0, 0, .., 0)$$

is an extreme point of S.

THEOREM 4.2. A nonvacuous constraint set of a linear program has at least one extreme point.

THEOREM 4.3. X_0 is an extreme point of a nonvacuous constraint set if and only if A can be decomposed into [D N], and X^T into $[X_D^T\ X_N^T]$ such that D is an mxm invertible matrix,

and

$$X_0 = \begin{bmatrix} X_D \\ X_N \end{bmatrix} = \begin{bmatrix} D^{-1} \\ \boldsymbol{O} \end{bmatrix} \geq \boldsymbol{O}.$$

Matrix D of Theorem 4.3 will be called a **Basis Matrix**, and the collection of all of its column vectors will be called a **Basis**. This theorem has an immediate result.

COROLLARY 4.3.1. The number of extreme points in the contraint set of a linear program is finite.

THEOREM 4.4. If a linear programming problem is optimal, then there exists an extreme point in the contraint set at which the objective function assumes the optimal value.

PROOF. By symmetry, it suffices to prove only for the maximal problem. By Corollary 4.3.1, let $X_1, X_2, .., X_t$ be all of the extreme points. From hypothesis, Let X_m be an optimal point which is not an extreme point. Then, we can write

$$X_m = \sum_{k=1}^{t} \lambda_k X_k, \quad \lambda_k' s \geq 0, \ \sum_{k=1}^{t} \lambda_k = 1.$$

It follows that

$$C^T X_m = C^T \sum_{k=1}^{t} \lambda_k X_k = \sum_{k=1}^{t} \lambda_k C^T X_k.$$

Since for some p = 1,2,..,t,

$$\text{Max } \{C^T X_k \mid k = 1,2,..,t\} = C^T X_p,$$

we have

$$C^T X_m \leq C^T X_p \sum_{k=1}^{t} \lambda_k = C^T X_p.$$

But, $C^T X_m$ is the maximal value of the objective function. We thus have

$$C^T X_m \geq C^T X_p.$$

It follows that

$$C^T X_m = C^T X_p.$$

This means that the objective value function assumes its maximal value at the extreme point X_p. ■

DEFINITION 4.6. (**Extreme Directions**) Let S be a nonvacuous closed convex set in E_n. A nonzero vector D in E_n is called a **Direction** of S if for each $X \in S$,

$$X + \lambda D \in S \text{ for all } \lambda \geq 0.$$

Two directions D_1, D_2 are called **Distinct** if

$$D_1 \neq \alpha D_2 \text{ for any } \alpha > 0.$$

A direction is called **Extreme** if it cannot be written as a positive linear combination of two distinct directions, i.e., if

$$D = \lambda_1 D_1 + \lambda_2 D_2 \text{ for } \lambda_1, \lambda_2 > 0,$$

then

$$D_1 = \alpha D_2 \text{ for some } \alpha > 0.$$

Clearly, a nonzero vector of a polyhedral convex cone is a direction, and it becomes extreme if it lies in an edge of the cone. Moreover, if S is a nonvacuous constraint set, then it follows from the structure of S that a vector $D \neq \emptyset$ is a direction of S if and only if

$$AD = \mathbf{0}, \text{ and } D \geq \mathbf{0}.$$

This characterization of a direction is useful in proving the next theorem that characterizes an extreme direction. However, as also for Theorem 4.7, we omit the proof, and direct interested readers to Bazaraa, Sherali and Shetty [1].

THEOREM 4.5. Let S be a nonvacuous constraint set of a linear program. A vector E is an extreme direction of S if and only if A can be decomposed into [D N] such that

$$D^{-1}A_j \leq \mathbf{0} \text{ for some } A_j \text{ of } N,$$

and

$$E \text{ is a positive multiple of } [(-D^{-1}A_j)^T\ \mathbf{e}_j^T]^T.$$

THEOREM 4.6. Let S be the nonvacuous constraint set of a linear program. Let $X_1, X_2, .., X_k$ be all of the extreme points, and $D_1, D_2, .., D_r$ be all of the extreme directions of S. Then, $X \in S$ if and only if it can be written as

$$X = \sum_{j=1}^{k} \lambda_j X_j + \sum_{j=1}^{r} \mu_j D_j,$$

where

$$\sum_{j=1}^{k} \lambda_j = 1,\ \lambda_j \geq 0 \ for\ j = 1,2,..,k,$$

and

$$\mu_j \geq 0 \ for\ j = 1,2,..,r.$$

We next state and prove a theorem that characterizes the optimality of a linear program.

THEOREM 4.7. Let $D_1, D_2, .., D_r$ be all of the extreme directions of the nonvacuous constraint set S of a minimal linear program. Then, this program has a finite optimal solution if and only if

$$C^T D_j \geq 0 \text{ for } j = 1,2,..,r.$$

PROOF. From Theorem 4.6, we write the given program in the following form:

Minimize

$$C^T\left(\sum_{j=1}^{k} \lambda_j X_j + \sum_{j=1}^{r} \mu_j D_j\right)$$

Subject to

$$\sum_{j=1}^{k} \lambda_j = 1,\ \lambda_j \geq 0 \ for\ j = 1,2,..,k,$$

and

$$\mu_j \geq 0 \ for \ j = 1,2,..,r.$$

If $C^T D_j < 0$ for some j, then the objective function can be made arbitrarily small by choosing μ_j large. This shows that a necessary and sufficient condition for a finite optimal solution is that

$$C^T D_j \geq 0 \text{ for } j = 1,2,..,r.$$

Furthermore, if this condition holds true, then, by choosing

$$\mu_j = 0 \text{ for } j = 1,2,..,r,$$

we can reduce the problem to the new one:

Minimize

$$C^T(\sum_{j=1}^{k} \lambda_j X_j)$$

subject to

$$\sum_{j=1}^{k} \lambda_j = 1; \ \lambda_j \geq 0 \ for \ j = 1,2,..,k.$$

Let

$$\lambda_i = 1, \text{ and } \lambda_j = 0 \text{ for } j \neq i,$$

where i is given by

$$C^T X_i = \inf \{C^T X_j | \ j = 1,2,..,k\}.$$

Then, an extreme point is chosen to optimize the problem. ■

The following theorem will set up a framework for an algorithm that can be used to solve a linear program in practice.

THEOREM 4.8. If X_* is an extreme point of the constraint set of a minimal linear program, then it is optimal or there exists a vector E such that

$$C^T(X_* + \lambda E) < C^T X_* \text{ for } \lambda > 0.$$

Moreover, E will be an extreme direction of X or there exists some $\lambda_* > 0$ such that $X_* + \lambda_* E$ is an extreme point.

PROOF. By Theorem 4.3, A can be decomposed into [D N], where $D = [D_1, D_2, \ldots, D_m]$ is an mxm matrix of full rank, N is an mx(n-m) matrix, and X_* can be written as

$$X_*^T = [X_{*B}^T \; X_{*N}^T] = [B'^T \; \mathbf{O}],$$

where

$$B' = D^{-1}B \geq \mathbf{O}.$$

Let $X \in S$. Then, it can be written as

$$X^T = [X_D^T \; X_N^T], \text{ where } X_D, X_N \geq \mathbf{O},$$

and

$$AX = [D \;\; N]\begin{bmatrix} X_D \\ X_N \end{bmatrix} = DX_D + NX_N = B.$$

Hence,

$$X_B = D^{-1}B - D^{-1}NX_N.$$

It follows that

$$C^{\top}X = \begin{bmatrix} C_D^{\top} & C_N^{\top} \end{bmatrix} \begin{bmatrix} X_D \\ X_N \end{bmatrix} = C_D^{\top}X_D + C_N^{\top}X_N$$

$$= C_D^{\top}D^{-1}B - C_D^{\top}D^{-1}NX_N + C_N^{\top}X_N$$

$$= C_D^{\top}D^{-1}B + (C_N^{\top} - C_D^{\top}D^{-1}N)X_N.$$

Therefore, if

$$C_N^{\top} - C_D^{\top}D^{-1}N \geq \mathbf{O},$$

then, since $X_N \geq 0$, we have

$$C^{\top}X \geq C^{\top}X_*,$$

and X is an optimal extreme point. Otherwise, suppose that

$$c_j - C_D^{\top}D^{-1}A_j < 0,$$

and consider

$$X = X_* + \lambda E_j, \text{ where } E_j = [(-D^{-1}A_j)^{\top}\ \mathbf{e}_j^{\top}]^{\top}.$$

Then,

$$C^{\top}X = C^{\top}X_* + \lambda C^{\top}E_j$$

$$= C^{\top}X_* + \lambda[C_D^{\top}\ C_N^{\top}][(-D^{-1}A_j)^{\top}\ \mathbf{e}_j^{\top}]^{\top}$$

$$= C^{\top}X_* + \lambda([C_D^{\top}(-D^{-1}A_j)+C_N^{\top}\mathbf{e}_j])$$

$$= C^{\top}X_* + \lambda(c_j-C_D^{\top}D^{-1}A_j). \tag{4.10}$$

Therefore, since $c_j-C_D^{\top}D^{-1}A_j < 0$, we have

$$C^{\top}X < C^{\top}X_* \text{ for } \lambda > 0.$$

We next consider two cases for $y_j = D^{-1}A_j$.

Case 1: $y_j \leq \mathbf{O}$. Then, $E_j \geq \mathbf{O}$. Consider

$$AE_j = [D\ N][(-D^{-1}A_j)^{\top}\ \mathbf{e}_j^{\top}]^{\top}$$

$$= -A_j + N\mathbf{e}_j = \mathbf{O} \text{ (since } N\mathbf{e}_j = A_j).$$

Since $AX_* = B$, for $X = X_* + \lambda E_j$, we have

$$AX = AX_* + \lambda AE_j = AX_* = B \text{ for all } \lambda.$$

Hence, X is feasible if and only if

$$X \geq \mathbf{O} \text{ for } E_j \geq \mathbf{O}.$$

It follows from (4.10) that the objective value function is unbounded. In this case, we have found an extreme direction E_j with

$$C^T E_j = c_j - C_D^T D^{-1} A_j < 0.$$

Case 2: $Y_j \nleq \mathbf{O}$. Let $B' = D^{-1}B$. Define

$$\lambda = \min \left\{ (B')_i / (Y_j)_i \,\middle|\, (Y_j)_i > 0 \right\} = (B')_r (Y_j)_r \geq 0, \qquad (4.11)$$

where $(B')_i$ and $(Y_j)_i$ denote the i-th component of B' and Y_j, respectively. For each basis element S, denote the element of X corresponding to S by x_S. Then, the components of $X_* + \lambda E_j$ are given by

$$x_{D_i} = B'_i - \frac{B'_r}{(Y_j)_r} (Y_j)_i \text{ for } i \in \{1, 2, \ldots, r-1, r+1, \ldots, m\}.$$

$$x_r = B'_r / (Y_j)_r, \qquad (4.12)$$

other x_i's are equal to zero.

Therefore,

$$x_{D_1}, x_{D_2}, \ldots, x_{D_{r-1}}, x_{D_{r+1}}, x_{D_m}, x_j$$

are the only positive components of X, so that X has at most m components to be positive. Moreover, The corresponding columns of A are linear independent. It follows from Theorem 4.3 that X is an extreme point. ■

DEFINITION 4.7. (**Simplex Algorithm**) The following procedure is called the **Simplex Algorithm**:

Initialzation step: Start with an extreme point X.

Main steps:

1. Write X in terms of D defined in Theorem 4.3, and calculate $C_D^T D^{-1} N - C_N^T$. If this vector is nonpositive, then stop; X is an optimal extreme point. Otherwise, pick the most positive $C_D^T D^{-1} A_j - c_j$. If $Y_j = D^{-1} A_j \le \mathbf{0}$, then stop; the objective value is unbounded along the the ray

$$\{X + \lambda[(-Y_j)^T\ \mathbf{e}_j^T]^T \mid \lambda \ge 0\}.$$

If $Y_j \not\le \mathbf{0}$, then go to step 2.

2. Compute the index r from (4.11) and form the new extreme point X in (4.12). Repeat step 1.

DEFINITION 4.8. (**Degeneracy**) If a basic feasible solution, defined in Definition 4.4, has exactly m positive entries, then it is called a **Nondegenerate** (**Basic Feasible**) Solution. Otherwise, it is said to be **Degenerate**. A linear program of which each basic feasible solution is not degenerate is called a **Nondegenerate Program.**

Clearly, a degenerate basic solution accompanies some coincidence of basic solutions. For example, if

$$A = \begin{bmatrix} 1 & 1 & 0 \\ 3 & 1 & 2 \end{bmatrix},\ B = \begin{bmatrix} 1 \\ 3 \end{bmatrix},$$

then, choosing x_1, x_2, or x_1, x_3 leads to the same basic solution, and hence degeneracy must occur. Conversely, a degenerate basic solution of AX = B, say $X = [1\ 0\ 0]^T$, can determine a coincidence of two basic solutions corresponding to basic variables x_1, x_2, or x_1, x_3 , respectively. For theoretical simplicity, nondegenerate programs will be assumed throughout this paper.

THEOREM 4.9. If the simplex algorithm is applied on a nondegenerate linear program, then it stops in a finite number of iterations.

DEFINITION 4.9. (Related to Theorem 4.3) If X is an extreme point, and D is the basis matrix corresponding to X, then the following block matrix (or its transformation) will be called a **Simplex Tableau**:

O	N	D	B
1	$-C_N^T$	$-C_D^T$	0

If T is transformed to T' by any elementary row operation on the upper submatrix [**O** N D B] or the row operation performed on the bottom row by adding to which a multiple of a higher row, then T' will be called an **Equivalent Simplex Tableau** of T. The corresponding linear programs to T and T' will be said to be **Equivalent programs**.

THEOREM 4.10. If

$$T = \begin{bmatrix} \mathbf{O} & N & D & B \\ 1 & -C_N^{\top} & -C_D^{\top} & u \end{bmatrix},$$

and

$$T' = \begin{bmatrix} \mathbf{O} & N' & D' & B' \\ 1 & -C'^{\top}_{N'} & -C'^{\top}_{D'} & u' \end{bmatrix}$$

are equivalent simplex tableaux, then the corresponding programs:

Minimize $\quad C^{\top}X + u \ \ (C^{\top} = [C_N^{\top}\ C_D^{\top}])$

subject to $\quad AX = B, \ \ (A = [N\ D]),$

and

Minimize $\quad C'^{\top}X' + u' \ \ (C'^{\top} = [C'^{\top}_{N'}\ C'^{\top}_{D'}])$

subject to $\quad A'X' = B', \ \ (A' = [N'\ D'])$

have the same constraint set, and the objective functions, restricted to the (common) constraint set, are actually identical.

PROOF. First, it clear that elementary row operations performed on $[\mathbf{O}\ N\ D\ B]$ will not change the constraint set. For the row operation performed on the bottom row, we have

$$[-C'^{\top}_{N'}\ -C'^{\top}_{D'} \mid u'] = [-C_N^{\top}\ -C_D^{\top} \mid u] + W^{\top}[N\ D \mid B]$$

for some $W \in E_m$. Note that

$$W^{\top}[N\ D \mid B] = [W^{\top}[N\ D] \mid W^{\top}B].$$

Thus, we have

$$-C'^{\top}_{N'}\ -C'^{\top}_{D'} = -C_N^{\top} + W^{\top}[N\ D \mid B],$$

and

$$u' = u + W^T B.$$

Let Z be in the common constraint set. Then, by defining

$$M = C^T Z + u, \text{ and } M' = C'^T Z + u',$$

we have that

$$\begin{aligned} M - M' &= (C^T - C'^T)Z + (u - u') \\ &= W^T [N\ D] Z - W^T B \\ &= W^T ([N\ D] Z - B) \\ &= \mathbf{O} \text{ (since } [N\ D] Z = B.) \end{aligned}$$

This implies that M and M' are really the same function. ■

Consequently, we immediately have the following result.

COROLLARY 4.10.1. Theorem 4.10 is still true if the nonnegative conditions are added to the constraints.

Indeed, by multiplying the constraint rows by D^{-1}, adding C_D^T times the new constraint rows to the objective row, we can identify the previous simplex tableau with the following one, where $B' = D^{-1}B$:

$\mathbf{O}$	$D^{-1}N$	$\mathbf{I}$	B'
1	$C_D^T D^{-1}N - C_N^T$	$\mathbf{O}$	$C_D^T B'$

DEFINITION 4.10. (**Tableau Simplex Algorithm**) The following procedure is called a **Tableau Simplex Algorithm**:

Initialization step: Start with an extreme point X, and the preceding simplex tableau.

Main steps:

1. If $C_D^T D^{-1} N - C_N^T \leq \mathbf{O}$, then stop; the current extreme point is optimal. Otherwise, pick a nonbasic variable with negative $C_D^T D^{-1} A_j - c_j$ (upon examining the objective row). If $D^{-1} A_j \leq \mathbf{O}$, then stop; the problem is unbounded. Otherwise, go to step 2.
2. Compute the index r from (4.11), and pivote $(Y_j)_r$ as follows:
   i) Divide the r-th row by $(Y_j)_r$.
   ii) Multiply the new r-th row by $-(Y_j)_i$, and add the result to the i-th constraint row for $i = 1,2,..,m$, $i \neq r$.
   iii) Multiply the new r-th row by $-(C_D^T D^{-1} A_j - c_j)$ and add the result to the objective row.

Repeat step 1.

EXAMPLE 4.2. Solve the following linear program:

$$\begin{array}{ll} \text{Minimize} & x_1 - 4x_2 \\ \text{subject to} & -x_1 + 2x_2 \leq 4, \\ & x_1 + x_2 \leq 3, \\ & x_1,\ x_2 \geq 0. \end{array}$$

SOLUTION. Using slack variables $x_3 \geq 0$, and $x_4 \geq 0$, we write the given problem in the following form:

$$\begin{array}{lll} \text{Minimize} & x_1 - 4x_2 & \\ \text{subject to} & -x_1 + 2x_2 + x_3 \qquad = & 4, \\ & x_1 + x_2 \qquad + x_4 = & 3, \\ & x_1, x_2, x_3, x_4 \geq 0. & \end{array}$$

Or, in terms of matrix notations,

$$\begin{array}{ll} \text{Minimize} & C^T X \\ \text{subject to} & AX = B, \\ & X \geq \mathbf{O}, \end{array}$$

where

$$C = [1\ -4\ 0\ 0]^T, \ B = [4\ 3]^T,$$

and

$$A = \begin{bmatrix} -1 & 2 & 1 & 0 \\ 1 & 1 & 0 & 1 \end{bmatrix}.$$

If

$$X^T = [X_N^T\ X_D^T] = [\mathbf{O}\ (D^{-1}B)^T] \geq \mathbf{O},$$

then

$$AX = B.$$

Therefore,

$$X = [0\ 0\ 4\ 3]^T$$

is an extreme point by Theorem 4.3.

Using X as a starting point, we have the starting simplex tableau:

0	-1	2	1	0	4
0	1	1	0	1	3
1	-1	4	0	0	0

Continuing applying the tableau simplex algorithm, we generate the following tableaux:

0	1/2	1	1/2	0	2
0	3/2	0	-1/2	1	1
1	1	0	-2	0	-8

0	0	1	1/3	1/3	7/3
0	1	0	-1/3	2/3	2/3
1	0	0	-5/3	-2/3	-26/3

Therefore, the optimal point is $[2/3\ \ 7/3]^T$, and the optimal value is -26/3. This result can be rechecked with that of Example 4.1.

CHAPTER V

DUAL LINEAR PROGRAMS

In this chapter, we will define a dual linear program from a given one, and then discuss certain basic theorems that will be used to prove the duality theorem and the existence theorem. Finally, we will develop the concept of complementary slackness to derive the primal-dual algorithm. The main references of this chapter are Cooper and Steinberg [2], Goldman and Tucker [7], Hadley [9], and Tucker [12].

DEFINITION 5.1. The following linear programs will be called **Dual Linear Programs**:

$$\begin{aligned} &\text{Maximize} && z = C^T X \\ &\text{subject to} && AX \leq B, \\ &&& X \geq \mathbf{O}, \end{aligned} \tag{5.1}$$

and

$$\begin{aligned} &\text{Minimize} && w = B^T U \\ &\text{subject to} && A^T U \geq C, \\ &&& U \geq \mathbf{O}. \end{aligned} \tag{5.2}$$

In other words, one of the forms will be called the **Dual** of the other, which is called the **Primal**.

Clearly, it follows from Definition 5.1 that the dual of the dual (of a primal) will be the primal.

Conventionally, we may choose the maximal problem to be a primal and the minimal one to be a dual. Furthermore, by using slack variables, we can verify that the following pair of linear programs is, in fact, the dual pair:

$$\begin{aligned} \text{Maximize} \quad & z = C^T X \\ \text{subject to} \quad & AX = B, \\ & X \geq \mathbf{0}, \end{aligned} \tag{5.3}$$

and

$$\begin{aligned} \text{Minimize} \quad & w = B^T U, \\ \text{subject to} \quad & A^T U \geq C. \end{aligned} \tag{5.4}$$

Note that U is free from nonnegative conditions. The later pair of linear programs is, in fact, an equivalent form of the former one. We will use the later one to derive the primal-dual algorithm, and use the other to derive basic theorems.

We first state two basic lemmas whose proofs can be found in Cooper and Steinberg [2].

LEMMA 5.1. If X and U are feasible solutions of the primal and the dual, respectively, then

$$C^T X \leq B^T U.$$

LEMMA 5.2. If X and U are feasible, and $C^T X = B^T U$, then they are optimal.

We note that a pair of dual linear programs is totally defined by fixed matrices A, B, C, which altogether define matrix K in Corollary 3.4.1. Moreover, corresponding to a solution $[U_0^T\ X_0^T\ t_0]^T$, we have U_0, X_0 as feasible solutions to the dual and the primal, respectively, provided that $t_0 = 1$. This situation can be extended, more generally, to the following theorem.

THEOREM 5.1. (Related to Corollary 3.4.1) If $t_0 > 0$, then there are optimal vectors X and U for the dual programs such that

$$C^TX = B^TU,\ U + B > AX,\ A^TU + X > C.$$

PROOF. Suppose $t_0 > 0$. Then,

$$[U^T\ X^T\ t] = 1/t_0[U^T_0\ X_0^T\ t_0] = [U^T_0/t_0\ X_0^T/t_0\ 1]$$

is a solution of the self dual defined by the following matrix

$$K' = t_0K = \begin{bmatrix} \boldsymbol{O} & t_0A & -t_0B \\ -t_0A^T & \boldsymbol{O} & t_0C \\ t_0B^T & -t_0C^T & \boldsymbol{O} \end{bmatrix}$$

such that all properties (3.31)-(3.34) of Corollary 3.4.1 remain the same. Hence, by (3.31) and (3.32) with $t_0 = 1$, we have U, X as feasible solutions. By (3.33), Lemma 5.1 and Lemma 5.2, U, X are optimal and $C^TX = B^TU$. By dividing (3.35) and (3.36) by t_0, we have

$$U + B > AX, \text{ and } A^TU + X > C. \blacksquare$$

EXAMPLE 5.1. (Related to Example 3.7) Using Theorem

5.1, find the optimal solution of the following linear program:

Maximize $\qquad -x$

subject to $\qquad x \leq 1,$

$\qquad x \geq 0.$

SOLUTION. The dual linear program is

Minimize $\qquad u$

subject to $\qquad u \geq -1,$

$\qquad u \geq 0.$

By the comment preceding Theorem 5.1, we have

$$K = \begin{bmatrix} 0 & -1 & 1 \\ 1 & 0 & 1 \\ -1 & -1 & 0 \end{bmatrix}.$$

Thus, by Example 3.7,

$$x_0 = 0, \ u_0 = 0, \ t_0 = 8$$

form a solution for the self dual system (3.27).

It follows from the proof of Theorem 5.1 that

$$x_* = x_0/8 = 0, \ u_* = u_0/8 = 0$$

are optimal solutions of the primal and the dual that defined by fixed matrices $A' = 8A$, $B' = 8B$ and $C' = 8C$. Noticing that optimality of one system will imply optimality of the other, we conclude that $x_0 = 0$, $u_0 = 0$ are optimal solutions of the dual and the primal that defined by fixed matrices A, B and C. Clearly, the optimal value is zero. Moreover, it can be checked that

$$u_0 + B > Ax_0, \text{ and } A^T u_0 + x_0 > C.$$

It should be noted that Theorem 5.1 offered a way to solve a linear program by reducing the dual programs to the self dual system of linear inequalities, and solving this self dual system by means of Corollary 3.4.1 and Theorem 5.1. However, as has been seen, even though this way is valuable in theory, it is unfavorable in practical computations. This situation can be overcome by using the simplex algorithm of Chapter IV, or other modified algorithms that will be presented next.

THEOREM 5.2. (Related to Corollary 3.4.1) Suppose that $t_0 = 0$. Then, the following statements are true:

a) At least one of the dual programs has no feasible vector.

b) If the maximal problem has a feasible vector, then the set of its feasible vectors is unbounded and $C^T X$ is unbounded on this set. (A dual statement can be established for the minimal problem.)

PROOF. To prove (a), suppose that X is a feasible vector for the primal. Then, by (ii), we have

$$A^T U_0 \geq \mathbf{O} \text{ (since } t_0 = 0 \text{ by hypothesis.)}$$

It follows that

$$U_0^T AX \geq \mathbf{O} \text{ (since } X \geq \mathbf{O}.\text{)}$$

Besides, by (3.34) of Corollary 3.4.1, with $t_0 = 0$, we have

$$B^T U_0 < C^T X_0.$$

Since X is feasible, $AX \le B$. This implies that

$$0 \le U_0^T AX \le U_0^T B < C^T X_0. \tag{5.7}$$

Now, suppose also that the dual has a feasible vector U. Then, by a similar argument, we have

$$0 \ge U^T AX_0 \ge C^T X_0.$$

But this contradicts (5.7), and thus (a) is proved.

To prove (b), for each $\lambda \ge 0$, consider

$$X' = X + \lambda X_0,$$

where X is a feasible solution of the primal and X_0 is as in the Corollary 3.4.1. Then, clearly, $X' \ge \mathbf{0}$. Hence, by (3.31) of this corollary, with $t_0 = 0$, we have

$$A(X + \lambda X_0) = AX + \lambda AX_0 \le AX \le B.$$

Hence, X' defines an infinite ray of feasible vectors, and this proves the first assertion of (b).

It follows from (5.7) that

$$C^T X_0 > 0.$$

Therefore, by choosing λ,

$$C^T (X + \lambda X_0) = C^T X + \lambda C^T X_0$$

can be made arbitrarily large. This proves the second assertion. ■

We immediately have the following Corollary.

COROLLARY 5.2.1 If one of the dual problems (and thus both) is optimal, then $t_0 > 0$.

We state next two theorems that characterize optimality in terms of boundedness, and duality. Interested readers may find the proofs in Cooper and Steinberg [2].

THEOREM 5.3. (**Boundedness Theorem**) A necessary and sufficient condition for either (and thus both) of the dual programs to have optimal vectors is that either C^TX or B^TU is bounded on the corresponding nonvacuous set of feasible vectors.

THEOREM 5.4. (**Duality Theorem**) A vector X_0 is an optimal solution for the primal if and only if there exists a feasible solution V_0 for the dual such that

$$B^TV_0 = C^TX_0.$$

(A similar dual statement can be obtained.)

For convenience of later use, we restate the duality theorem in another version.

COROLLARY 5.4.1. The primal is optimal if and only if the dual is optimal. Moreover, If X_0, and U_0 are optimal solutions for the primal and the dual, respectively, then we have

$$C^TX_0 = B^TU_0.$$

Before presenting the existence theorem, we give an example to illustrate the duality theorem.

EXAMPLE 5.2. Solve the following program by using its dual:

Minimize $\quad -x_1 + 4x_2$

Subject to

$$
\begin{aligned}
x_1 + x_2 &\le 3,\\
-x_1 + 2x_2 &\le 4,\\
x_1 &\le 2,\\
-x_2 &\le -1,\\
x_1, x_2 &\ge 0.
\end{aligned}
$$

SOLUTION. The dual program is given by

Maximize $\quad 3u_1 + 4u_2 + 2u_3 - u_4$

Subject to

$$
\begin{aligned}
u_1 - u_2 + u_3 &\ge -1,\\
u_1 + 2u_2 - u_4 &\ge 4,\\
u_1, u_2, u_2, u_4 &\ge 0.
\end{aligned}
$$

We apply the tableau simplex algorithm, using surplus variables, to generate the following simplex tableaux:

1	-1	1	0	-1	0	-1
1	2	0	-1	0	-1	4
-3	-4	-2	1	0	0	0

3/2	0	1	-1/2	-1	-1/2	1
1/2	1	0	-1/2	0	-1/2	2
-3	-4	-2	1	0	0	0

3/2	0	1	-1/2	-1	-1/2	1
1/2	1	0	-1/2	0	-1/2	2
2	0	0	-2	-2	-3	10

1	0	2/3	-1/3	-2/3	-1/3	2/3
0	1	-1/3	-1/3	1/3	-1/3	5/3
0	0	-4/3	-4/3	-2/3	-7/3	26/3

The optimal point and optimal value for the dual are

$U = [2/3\ 5/3\ 0\ 0]^T$, and $w = 26/3$.

By Theorem 5.4, we have

$$X = (C_D^T D^{-1})^T,$$

where $C_D^T = [3\ 4]$, and

$$D = \begin{bmatrix} 1 & -1 \\ 1 & 2 \end{bmatrix}.$$

Hence,

$$D^{-1} = \begin{bmatrix} \frac{2}{3} & \frac{1}{3} \\ -\frac{1}{3} & \frac{1}{3} \end{bmatrix},$$

and

$$X = [2/3 \ \ 7/3]^T.$$

The next theorem characterizes optimality as feasibility of both programs.

THEOREM 5.5. (**Existence Theorem**) A linear program has a finite optimal solution (and so does its dual) if and only if it and its dual are feasible.

We next gives two examples, illustrating circumstances in which at least one of the dual programs has a nonempty constraint set, i.e., nonfeasible, and hence each of the programs must be nonoptimal by the existence theorem.

EXAMPLE 5.3. Consider the following problem:

Maximize $\quad -x_1 - x_2$

Subject to $\quad -x_1 - x_2 \le -2,$

$x_1 + 2x_2 \le 1,$

$x_1, x_2 \ge 0.$

Since $x_2 \le -1$, and $x_2 \ge 0$, the primal is not feasible.

Next, consider its dual program that is given by

$$\begin{array}{lrcl} \text{Minimize} & -2u_1 + u_2 & & \\ \text{Subject to} & -u_1 + u_2 & \geq & -1, \\ & -u_1 + 2u_2 & \geq & -1, \\ & u_1, u_2 & \geq & 0. \end{array}$$

Then, $U = [0\ \ 0]^T$ meets the constraints. Hence, the dual is feasible. However, it is not optimal by the existence theorem.

EXAMPLE 5.4. Consider the following problem:

$$\begin{array}{lrcl} \text{Maximize} & -x_1 + 2x_2 & & \\ \text{Subject to} & 0x_1 + 2x_2 & \leq & -2, \\ & 0x_1 - 1x_2 & \leq & 1, \\ & x_1, x_2 & \geq & 0. \end{array}$$

Since $x_2 \leq -1$, and $x_2 \geq 0$, the primal is not feasible.

Next, consider its dual program that is given by

$$\begin{array}{lrcl} \text{Minimize} & -2u_1 + u_2 & & \\ \text{Subject to} & 0u_1 + 0u_2 & \geq & 1, \\ & 2u_1 - u_2 & \geq & 2, \\ & u_1, u_2 & \geq & 0. \end{array}$$

Then, it is also not feasible.

To illustrate how the theorem works at another case, we give the following example.

EXAMPLE 5.5 Consider the problem:

$$\begin{array}{lrcl} \text{Maximize} & -x_1 + 2x_2 & & \\ \text{Subject to} & 0x_1 + 2x_2 & \leq & -2, \end{array}$$

$$0x_1 - 1x_2 \leq 1,$$

$$x_1, x_2 \geq 0.$$

Then, it is clear that $[3\ 1]^T$ is a feasible solution.

Now, consider the dual linear program:

Minimize $\quad 2u_1 - u_2$

Subject to $\quad u_1 - u_2 \geq 1,$

$$-u_1 + 2u_2 \geq 2.$$

$$u_1, u_2 \geq 0.$$

Then, $[4\ 3]^T$ is clearly a feasible solution. So, by the existence theorem, both programs are optimal. (The common optimal value can be immediately obtained.)

We now develop the so-called two-phase tableau simplex algorithm which is more effective than the tableau simplex method in the sense that it can be used to find an initial feasible solution and then use this solution to obtain the optimal solution. This algorithm, in fact, has features on which the primal-dual algorithm is developed.

DEFINITION 5.2. For each maximal linear program of the Form (4.4)-(4.6), the following program:

Maximize $\quad z_a = -\mathbf{1}^T X_a$

subject to $\quad AX + X_a = B, \qquad (5.9)$

$$[X\ X_a] \geq \mathbf{O}$$

is called the **Artificial Linear Program** generated by the given maximal one. X_a is called the vector of **Artificial variables**.

From (4.4)-(4.6), we may assume $B \geq \mathbf{0}$.

DEFINITION 5.3. (**Two-Phase Tableau Simplex Algorithm**) The following procedure is called the **Two-Phase Tableau Simplex Algorithm**:

Phase I: Solve the artificial problem (5.9), using the tableau simplex algorithm (Definition 4.10). Note that since the artificial problem is bounded above by 0, it is always optimal. Furthermore, its initial feasible solution can be immediately obtained by treating artificial variables as basic ones. If the optimal value $\neq 0$, then stop; the original problem is unbounded. (Otherwise, go to Phase II.)

Phase II: Apply the tableau simplex algorithm again for the original problem, using the optimal point of Phase I as an initial point.

EXAMPLE 5.6. Solve the following program:

$$\begin{array}{ll} \text{Maximize} & -x_1 + 4x_2 \\ \text{Subject to} & x_1 + 4x_2 + x_3 + 2x_4 = 10, \\ & 2x_1 + 5x_2 + x_3 + x_4 = 14. \end{array}$$

SOLUTION. Let x_5, x_6 be the artificial variables. Then, an initial solution for the artificial problem is

$$x_1 = x_2 = x_3 = x_4 = 0,\ x_5 = 10,\ x_6 = 14.$$

Phase I: By applying the tableau simplex algorithm, we generate the simplex tableaux for the artificial problem:

0	1	4	1	2	1	0	10
0	2	5	1	1	0	1	14
1	0	0	0	0	1	1	0

0	1	4	1	2	1	0	10
0	2	5	1	1	0	1	14
1	-3	-9	-2	-3	0	0	-24

0	1/4	1	1/4	1/2	1/4	0	5/2
0	3/4	0	-1/4	-3/2	-5/4	1	3/2
1	3/4	0	1/4	3/2	9/4	0	-3/2

0	0	1	1/3	1	2/3	-1/3	2
0	1	0	-1/3	-2	-5/3	4/3	2
1	0	0	0	0	1	1	0

Each entry of the objective row is nonnegative. So, Phase I is ended with an optimal point X given by

$$X = [2\ 2\ 0\ 0\ 0\ 0\ ]^T,$$

and the optimal value is

$$Z_a = 0.$$

Phase II: We use the point $X = [2\ 2\ 0\ 0]^T$ as an initial extreme point for the original problem. Again, by applying the tableau simplex algorithm for the original problem, we generate the simplex tableaux:

0	1	4	1	2	10
0	2	5	1	1	14
1	1	-4	0	0	0

0	1	0	-1/3	-2	2
0	0	1	1/3	1	2
1	0	0	5/3	6	6

Since each entry of the objective row is nonnegative, the original problem is optimized with the optimal point $[2\ 2\ 0\ 0]^T$, and the optimal value is 6.

The next lemma and its corollary are important in establishment of the equilibrium theorem.

LEMMA 5.3. The self dual, defined in Definition 3.2, possesses a solution W' such that each of the pairs

$$\sum_{j=1}^{n} k_{ij}w_j \geq \boldsymbol{O}, \quad w_j \geq \boldsymbol{O} \ (j=1,2,..,n), \tag{5.10}$$

contains exactly one inequality that is strict with respect to W'.

PROOF. For any solution W, we have

$$W^T KW = \sum_{i=1}^{n} w_i (\sum_{j=1}^{n} k_{ij}w_j) = \sum_{i=1}^{n} (\sum_{j=1}^{n} w_i k_{ij} w_j)$$

$$= \sum_{i=1}^{n} (\sum_{j=1}^{n} w_i (-k_{ji}) w_j) = - \sum_{i=1}^{n} (\sum_{j=1}^{n} w_j k_{ji} w_i)$$

$$= - \sum_{j=1}^{n} (\sum_{i=1}^{n} w_j k_{ji} w_i) = -W^T KW.$$

Hence,

$$\boldsymbol{O} = W^T KW = \sum_{i=1}^{n} w_i (\sum_{j=1}^{n} k_{ij} w_j) .$$

It follows that each pair of (5.10) must have at least one inequality which is an equality. This implies that at most one of the pair can be strict with respect to some solution.

On the other hand, by Theorem 3.4, there exists a solution W' such that

$$\sum_{j=1}^{n} k_{ij}w_j' + w_i' > 0 \text{ for each } i = 1,2,..,n.$$

That is, one of the pair (5.10) must be strict with respect to W', and therefore, exactly one of the pair will be slack. ∎

The preceding lemma immediately has the following result.

COROLLARY 5.3.A. Let K be defined as in Corollary 3.4.1. Then, the self dual system, defined by Definition 3.2, possesses a solution $W' = [U'^{T}\ X'^{T}\ t]^{T}$ such that exactly one inequality from each following pair must be strict with respect to W':

$$\left(\sum_{j=1}^{m} k_{ij}u_j + \sum_{j=m+1}^{m+n} k_{ij}x_j + k_{i,m+n+1}t \geq 0,\ w_i \geq 0\right),$$

where

$$w_i = \left[\begin{array}{l} u_i \text{ if } i = 1,2,..,m \\ x_i \text{ if } i = m+1,..,m+n \\ t \text{ if } i = m+n+1 \end{array}\right).$$

In general, we say that a linear inequality is **Slack** if it is strict with respect to some solution, and when this is the case we also say it is slack with respect to this solution.

THEOREM 5.6. Consider a pair of dual programs defined in Definition 5.1. Then, for each pair of optimal solutions to this dual linear programming problem, say (X,X_s) and (U,U_s), we have that

$$x_j u_{sj} = 0 \text{ for each } j = 1,2,..,n,$$

and (5.11)

$$u_i x_{si} = 0 \text{ for each } i = 1,2,..,m.$$

PROOF. By adding slack variables and surplus variables, respectively for the primal and the dual, we can reduce the given dual problem to the following equivalent one:

$$\begin{aligned} &\text{Maximize} && z = C^T X \\ &\text{subject to} && AX + X_s = B, \\ & && X,\ X_s \geq \mathbf{O}, \end{aligned} \tag{5.12}$$

and

$$\begin{aligned} &\text{Minimize} && w = B^T U \\ &\text{subject to} && A^T U - U_s = C, \\ & && U,\ U_s \geq \mathbf{O}. \end{aligned} \tag{5.13}$$

Multiplying (5.12) by U^T, we have

$$U^T AX + U^T X_s = U^T B.$$

Taking the transpose for both sides,

$$X^T A^T U + X_s^T U = B^T U. \tag{5.14}$$

Similarly, multiplying (5.13) by X^T, we get

$$X^T A^T U - X^T U_s = X^T C = C^T X. \tag{5.15}$$

Since (X,X_s) and (U,U_s) are optimal, by Corollary 5.4.1, we have

$$C^TX = B^TU. \tag{5.16}$$

Therefore, adding up (5.14) and the negative of (5.15), it follows from (5.16) that

$$U^TX_s + X^TU_s = 0. \tag{5.17}$$

Since X, U, X_s, U_s are all ncnnegative, we have

$$U^TX_s = 0, \; X^TU_s = 0. \tag{5.18}$$

That is,

$$x_ju_{sj} = 0 \text{ for each } j = 1,2,..,n,$$

and (5.19)

$$u_ix_{si} = 0 \text{ for each } i = 1,2,..,m. \quad \blacksquare$$

EXAMPLE 5.7. Consider the the pair of dual linear programs given in Example 5.2. Then, we have

$$X = [2/3 \;\; 7/3]^T, \; X_s = [0 \;\; 0 \;\; 4/3 \;\; 4/3]^T,$$

and

$$U = [2/3 \;\; 5/3 \;\; 0 \;\; 0]^T, \; U_s = [0 \;\; 0]^T.$$

Clearly,

$$U^TX_s = 0, \text{ and } X^TU_s = 0,$$

which verify (5.18), and (5.19).

DEFINITION 5.4. Property (5.18), or (5.19), is called **Complementary Slackness** (or **Equilibrium**).

The following alternative proof for Theorem 5.6 is essentially obtained by using Lemma 5.3 for matrix K defined in Corollary 3.4.1.

ALTERNATIVE PROOF. (For Theorem 5.6) Since the given dual linear programs are optimal, $t_0 > 0$ by Corollary 5.2.1. By applying Corollary 5.3.A, we have the following cases.

Case 1: $i = 1,2,..,m$. Then, $w_i = u_i$. If $u_i = 0$, then $u_i x_{si} = 0$. So, suppose $u_i > 0$. Then, by Corollary 5.3.A, we have

$$\sum_{j=1}^{m} k_{ij}u_j + \sum_{j=m+1}^{m+n} k_{ij}x_j + k_{i,m+n+1}t = 0.$$

Hence, since $k_{ij} = 0$ for $i \in \{1,2,..,m\}$, we have

$$\sum_{j=m+1}^{m+n} k_{ij}x_j + k_{i,m+n+1}t = 0.$$

Denote the k-th column vector of a general matrix, say M, by M_k. Then, we can write

$$(-A^T)_i{}^T X + b_i t = 0.$$

That is,

$$(A^T)_i{}^T X = b_i t.$$

From the proof of Theorem 5.1, without loss of generality, we may assume that $t = 1$. Hence, we have

$$(A^T)_i{}^T X = b_i. \tag{5.21}$$

Since X is feasible (in fact, it is optimal), we can compare (5.21) with the form

$$(A^T)_i{}^T X + X_{si} = b_i,$$

so that

$$X_{si} = 0 \text{ for } i = 1,2,..,m.$$

Case 2: $i = m+1,m+2,..,m+n$. With a similar argument as in case 1, we can obtain

$$u_{si} = 0 \text{ for } i = m+1, m+2, .., m+n,$$

provided that $x_i > 0$. ■

As the existence theorem characterizes optimality by the existence of a pair of feasible solutions, the following theorem together with Theorem 5.6 will characterize a specific pair of optimal vectors. As will be seen later, it plays an important role in establishment of the primal-dual algorithm.

THEOREM 5.7. The feasible solutions (X, X_s), and (U, U_s) of the primal and the dual, respectively, are optimal if the complementary slackness holds.

PROOF. Suppose that complementary slackness holds. Then, we have

$$U^T X_s + X^T U_s = 0.$$

Hence,

$$U^T X_s = -X^T U_s = -U_s^T X.$$

Adding $U^T AX$ to both sides of this equation, we have

$$U^T(AX + X_s) = (U^T A - U_s^T)X = X^T(A^T U - U_s).$$

Since

$$AX + X_s = B, \text{ and } A^T U - U_s = C,$$

we have

$$U^T B = X^T C.$$

That is,

$$C^TX = B^TU.$$

Hence, by Lemma 5.2, X and U are optimal. ■

Immediately, we have the following Corollary.

COROLLARY 5.7.1. Solutions X and U of the primal and the dual, of the form (5.3), (5.4), respectively, are optimal if and only if

$$X^TU_s = 0, \tag{5.22}$$

where U_s is the solution of the dual slack variables.

EXAMPLE 5.8. Consider the pair of dual linear programs:

$$\begin{array}{ll} \text{Maximize} & -x_1 + 4x_2 \\ \text{Subject to} & x_1 + x_2 + x_3 = 3, \\ & -x_1 + 2x_2 + x_4 = 4, \\ & x_1 + x_5 = 2, \\ & -x_2 + x_6 = -1, \\ & x_1,\ x_2,\ x_3,\ x_4,\ x_5,\ x_6 \ge 0, \end{array}$$

and

$$\begin{array}{ll} \text{Minimize} & 3u_1 + 4u_2 + 2u_3 - u_4 \\ \text{Subject to} & u_1 - u_2 + u_3 \ge -1, \\ & u_1 + 2u_2 - u_4 \ge 4, \\ & u_1,\ u_2,\ u_3,\ u_4 \ge 0. \end{array}$$

First, it can be verified that

$X = [2/3\ \ 7/3\ \ 0\ \ 0\ \ 4/3\ \ 4/3]^T$, and $U = [2/3\ \ 5/3\ \ 0\ \ 0]^T$

are feasible solutions of the primal and the dual. This implies that

$$U_s = [0\ 0\ 2/3\ 5/3\ 0\ 0]^T.$$

Therefore,

$$X^T U_s = \mathbf{0}.$$

By Corollary 5.7.1, we conclude that X and U are optimal solutions for the primal and the dual problem, respectively.

DEFINITION 5.5. Without loss of consistency, we also call (5.22) **Complimentary Slackness** (or **Equilibrium**).

The next development aims to construct the so-called primal-dual algorithm basing the ideas on complimentary slackness. In particular, Corollary 5.7.1 will be used to test optimality for the problem in this algorithm.

The two-phase tableau simplex algorithm actually generates a finite sequence of extreme points of the artificial linear program (5.9) after Phase I is terminated. Furthermore, the last extreme point of this sequence, say

$$X = (x_1, x_2, .., x_n, x_{n+1}, x_{n+2}, .., x_{n+m}),$$

has its first n entries forming an extreme point, say $(x_1, x_2, .., x_n)$, for the original problem (4.4)-(4.6). As a result, this extreme point may be far away from the optimal point of the original problem when Phase II starts.

To improve the situation, we rely on the dual problem, which is carrying vector C in the constraints, to select a vector (to enter a basis) that guarantees complimentary slackness for a dual solution and the first n entries of the new extreme solution. Therefore, an initial dual solution is really required to implement the ideas. Since the initial solution for the artificial program has its first n components to be zeros, we must have complimentary slackness for the first n components and the initial dual solution. Hence, the next performance is to choose vectors to enter and to leave the basis so that the next tableau still maintains complimentary slackness.

By (5.22), we actually need part of the dual solution that associates with n slack variables. From this, we define

$$W = \{A_j \mid V_{sj} = 0\}, \tag{5.23}$$

where A_j's denote column vectors of A. Then, by allowing only A_j's to enter the basis, we will preserve complimentary slackness. For if $V_{sj} > 0$, then the corresponding x_j cannot be a basic variable, and hence remains zero.

To formally develop the algorithm, we first state the following theorem whose proof is actually contained in Theorem 5.10, developed later.

THEOREM 5.8. Let W be defined as in (5.23), D be the basis obtained when the optimality criterion is satisfied for all vectors in W, and $\check{C}$ be the vector of prices of the artificial program (5.9). Then, there exists a $\theta > 0$ such that

$$V' = V + \theta \check{C}_D D^{-1} \qquad (5.24)$$

is a new solution to the dual, and the corresponding objective value

$$w' < w,$$

where V, and $\check{C}_D$ are respectively the dual solution, and the vector of prices in $\check{C}$ corresponding to D.

Moreover, the dual will be bounded (and hence optimal) if and only if the set of all θ's that hold for (5.24) is bounded. If this is the case, we have

$$v'_{sj} = v_{sj} + \theta_0(\check{z}_j - \check{c}_j), \qquad (5.25)$$

where

$$\check{z}_j = \check{C}_D D^{-1} A_j,$$

and

$$\theta_0 = \min \left\{ -v_{sj}/(\check{z}_j - \check{c}_j) \mid \check{z}_j - \check{c}_j < 0 \text{ and } v_{sj} > 0 \right\}. \qquad (5.26)$$

Otherwise, we have

$$\check{z}_j - \check{c}_j \geq 0 \text{ for } j = 1,2,..,n, \text{ and } z_a < 0. \qquad (5.27)$$

This theorem and the preceding discussion can be made in effect by means of the following definition.

DEFINITION 5.6. (Preliminary Primal-Dual Algorithm) The following procedure will be called The **Preliminary Primal-Dual Algorithm**:

Initialization Step: Start with an initial (not necessarily feasible) point of the dual constraints (5.4), for which $W = \{A_j \mid v_{sj} = 0\} \neq \emptyset$.

Main steps:

1. If $\check{z}_j - \check{c}_j \geq 0$ for $j = 1, 2, .., n$, then stop; the original program is either optimal (if $z_a = 0$) or unbounded (if $z_a < 0$).

Otherwise, if $\check{z}_j - \check{c}_j \geq 0$ for all $A_j \in W$, then compute θ_0 by (5.26), V_s by (5.25), and then update V_s, and W. (Otherwise, go to the next step.)

2. Drive a vector A_j in W with $\check{z}_j - \check{c}_j < 0$ into D, drive a corresponding vector out of D, and update the tableau in the same way done in Phase I of the two phase algorithm. Repeat step 1.

EXAMPLE 5.9. Solve the following program, using the preceding algorithm, assuming initially $V_s = [0\ 7\ 4\ 3]^T$:

$$
\begin{array}{lrcl}
\text{Maximize} & -x_1 + 4x_2 & & \\
\text{Subject to} & -x_1 + 2x_2 + x_3 \qquad & = & 4, \\
 & x_1 + x_2 \qquad + x_4 & = & 3, \\
 & x_1,\ x_2,\ x_3,\ x_4 & \geq & 0.
\end{array}
$$

SOLUTION. Applying the preliminary primal-dual algorithm, we generate the following simplex tableaux:

TABLEAU I			-1	4	0	0
$\check{C}_D$	X_D	Basis	A_1	A_2	A_3	A_4
-1	4	$\mathbf{e}_1$	-1	2	1	0
-1	3	$\mathbf{e}_2$	1	1	0	1
z_a	-7	$\check{z}_j - \check{c}_j$	0	-3	-1	-1
w	-	v_{sj}	0	7	4	3

TABLEAU II			-1	4	0	0
$\check{C}_D$	X_D	Basis	A_1	A_2	A_3	A_4
0	2	A_2	-1/2	1	1/2	0
-1	1	$\mathbf{e}_2$	3/2	0	-1/2	1
z_a	-1	$\check{z}_j - \check{c}_j$	-3/2	0	1/2	-1
w	-	v_{sj}	0	0	5/3	2/3

TABLEAU III			-1	4	0	0
$\check{C}_D$	X_D	Basis	A_1	A_2	A_3	A_4
0	7/3	A_2	0	1	1/3	1/3
0	2/3	A_1	1	0	-1/3	2/3
z_a	0	$\check{z}_j-\check{c}_j$	0	0	0	0
w	26/3	v_{sj}	0	0	5/3	2/3

Since $z_j - c_j = 0$, and $z_a = 0$, we obtain an optimal point $X = [2/3\ 7/3\ 0\ 0]^T$, and the optimal value

$$z = -1(2/3) + 4(7/3) = 26/3.$$

We now derive the primal-dual algorithm by improving the ideas developed for the preliminary primal-dual algorithm.

Again, we consider a linear program:

$$\begin{array}{lll} \text{Maximize} & z = C^T X & \\ \text{subject to} & AX = B, & (5.28) \\ & X \geq \mathbf{O}, & \end{array}$$

where C, A, B are nx1, mxn, mx1 matrices. Add a new constraint

$$x_0 + \textstyle\sum_{1\leq j\leq n} x_j = b_0, \qquad (5.29)$$

where $x_0 \geq 0$ and b_0 is a positive real number that is large enough so that (5.29) places no restriction on X.

We then modify the primal into the following one:

Maximize $\qquad z = C^T X$

subject to

$$\begin{bmatrix} 1 & \mathbf{1}^T \\ \boldsymbol{O} & A \end{bmatrix} \begin{bmatrix} x_0 \\ X \end{bmatrix} = \begin{bmatrix} b_0 \\ B \end{bmatrix}, \tag{5.30}$$

$$[x_0 \; X^T]^T \geq \mathbf{O}.$$

This modified primal problem has its dual in the form:

Minimize $\qquad w = b_0 v_0 + B^T V$

subject to

$$\begin{bmatrix} 1 & \boldsymbol{O} \\ \mathbf{1} & A^T \end{bmatrix} \begin{bmatrix} v_0 \\ V \end{bmatrix} \geq \begin{bmatrix} 0 \\ C \end{bmatrix}. \tag{5.31}$$

By defining $c_0 = 0$, the preceding dual problem immediately has an initial feasible solution:

$$v_0 = \max \{c_i\}, \text{ and } V = \mathbf{O}. \tag{5.32}$$

Substracting surplus variables from (5.31), we obtain

Minimize $\qquad w = B^T V$

subject to

$$\begin{bmatrix} 1 & \boldsymbol{O} \\ \mathbf{1} & A^T \end{bmatrix} \begin{bmatrix} v_0 \\ V \end{bmatrix} - \begin{bmatrix} 1 & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{I} \end{bmatrix} \begin{bmatrix} v_{s0} \\ V_s \end{bmatrix} = \begin{bmatrix} 0 \\ C \end{bmatrix}, \tag{5.33}$$

$$[x_{s0} \; V_s^T]^T \geq \mathbf{O}.$$

If $v_0 = 0$, then we obtain

Minimize $\qquad w = b_0 v_0 + B^T V$

subject to $\qquad A^T V - V_s = C, \qquad (5.34)$

$\qquad (\text{or } A^T V \geq C.)$

This newly defined linear program is actually the dual of the primal defined by (5.28), and any solution of (5.33) with $v_0 = 0$ can be considered as a solution for this dual problem (5.34). Moreover, by Corollary 5.7.1, if $[x_0\ X^T]^T$ is a feasible solution that preserves complimentary slackness, i.e.,

$$x_0 v_{s0} + X^T V_s = 0,$$

then $[x_0\ X^T]^T$, and $[v_0\ V^T]^T$ are optimal solutions to the modified problems (5.30), and (5.31). Hence, if $v_0 = 0$, then

$$z = C^T X = B^T V = w,$$

which is the optimal value of the dual problems (5.28), (5.34). Otherwise, if $v_0 \neq 0$, then

$$z = w = b_0 v_0 + B^T V,$$

which is the optimal value of the dual problems (5.30), (5.33).

The preceding discussion has actually proved the following theorem.

THEOREM 5.9. If $[x_0\ X^T]^T$ and $[v_0\ V^T]^T$ are respectively feasible solutions to the modified dual programs (5.30) and (5.33), then X and V are respectively optimal solutions to the original problems (5.28) and (5.34) if and only if

$$v_0 = 0 \text{ and } x_0 v_{s0} + X^T V_s = 0. \qquad (5.35)$$

From this theorem, we will use the same technique used for the preliminary primal-dual algorithm to develop an algorithm for the modified linear program (5.30), having (5.32) as an initial solution.

By adding artificial variables for the modified problem (5.30), we obtain the modified artificial problem:

Maximize $$z_a = -x_{a0} - \mathbf{1}^T X_a$$

subject to

$$\begin{bmatrix} 1 & \mathbf{1}^T \\ 0 & A \end{bmatrix} \begin{bmatrix} x_0 \\ X \end{bmatrix} + \begin{bmatrix} 1 & \mathbf{O} \\ \mathbf{O} & \mathbf{I} \end{bmatrix} \begin{bmatrix} x_{a0} \\ X_a \end{bmatrix} = \begin{bmatrix} b_0 \\ B \end{bmatrix}, \tag{5.36}$$

$$[x_0 \ X^T \ x_{a0} \ X_a^T] \geq \mathbf{O}.$$

Let α_j, j = 0,1,2,..,n be the j-th column of the first matrix, D be a basis matrix, $\check{C}$ be the the vector of prices, and $\check{C}_D$ as the vector of prices corresponding to basic variables. Initially, D will be chosen to consist entirely of artificial vectors. Besides, from (5.33), we have

$$\mathbf{1}v_0 + A^T V - V_s = C.$$

Therefore, Initially, since V = $\mathbf{O}$, we have

$$V_s = v_0 \mathbf{1} - C. \tag{5.37}$$

This implies that at least one component of V_s vanishes (since $v_0 = \max \{c_j\}$.)

What can be seen from the artificial problem (5.36) is that any feasible solution $[x_0 \ X^T \ 0 \ \mathbf{O}]^T$, once obtained, will maximize it with the maximal value of zero. Also, it supplies $[x_0 \ X^T]^T$ as a feasible solution for the modified primal problem (5.30). This situation motivates the idea of maximizing the artificial problem (5.36), and maintaining complimentary slackness at the same time. For when $z_a = 0$, we will find a feasible (and hence optimal) solution to the modified problem (5.30), whose optimal value is

$$z = w = b_0v_0 + B^TV.$$

In particular, we want to use the dual solution as a guide to select the vectors to enter the basis D at each iteration. From (5.37), we defined

$$W = \{\alpha_j \mid v_{sj} = 0\}. \tag{5.38}$$

It is clear that, initially, $W \neq \varnothing$. Therefore, by allowing only α_j's of W to enter the basis at any iteration, we maintain complimentary slackness. The vector to leave the basis and the updated tableau can be obtained in the same way performed in the simplex method. We continue iterating until

$$\check{z}_j - \check{c}_j \geq 0 \text{ for each } \alpha_j \text{ in } W, \tag{5.39}$$

where

$$\check{z}_j = \check{C}_D D^{-1} A_j.$$

When this condition occurs, we can still improve the dual solution by using the following theorem.

THEOREM 5.10. Let D be the basis in terms of condition (5.39). Then, there exists a $\theta > 0$ such that

$$[v'_0 \; V'^T] = [v_0 \; V^T] + \theta \check{C}_D D^{-1} \tag{5.40}$$

is a (new) solution to the dual and the corresponding objective value $w' < w$.

Moreover, the dual is bounded if and only if

$$\theta = \theta_0,$$

where

$$\theta_0 = \inf \{v_{sj}/(\check{z}_j - \check{c}_j) \mid \check{z}_j - \check{c}_j < 0\} < 0 \text{ and } v_{sj} > 0\} \tag{5.41}$$

$$< \infty.$$

In this case,

$$v'_{sj} = v_{sj} + \theta_0(\check{z}_j - \check{c}_j). \tag{5.42}$$

Otherwise, we have

$$\check{z}_j - \check{c}_j \geq 0 \text{ for } j = 1,2,..,n, \text{ and } z_a < 0. \tag{5.43}$$

PROOF. Since

$$V_s = v_0\mathbf{1} + C,$$

initially, we have $W \neq \varnothing$. For each $\theta \geq 0$, consider

$$[v'_0\ V'^{\top}] = [v_0\ V^{\top}] + \theta\check{C}_D D^{-1}. \tag{5.44}$$

Then, by multiplying both sides by $\alpha_j \in W$, we have

$$[v'_0\ V'^{\top}]\alpha_j = [v_0\ V^{\top}]\alpha_j + \theta\check{C}_D D^{-1}\alpha_j. \tag{5.45}$$

But

$$[v_0\ V^{\top}]\alpha_j - v_{sj} = \check{c}_j, \text{ and } v_{sj} = 0.$$

Thus,

$$\alpha_j^{\top}[v_0\ V^{\top}]^{\top} = \check{c}_j \text{ for each } \alpha_j \in W. \tag{5.46}$$

In addition,

$$\check{c}_j = 0 \text{ for each } j = 0,1,2,..,n.$$

It follows that

$$\check{z}_j - \check{c}_j = \check{z}_j = \check{C}_D D^{-1}\alpha_j. \tag{5.47}$$

Combining (5.45)-(5.47), we obtain

$$[v'_0\ V'^{\top}]\alpha_j = \check{c}_j + \theta(\check{z}_j - \check{c}_j) \text{ for } \alpha_j \in W.$$

So, by (5.39), we have

$$[v'_0\ V'^{\top}]\alpha_j \geq \check{c}_j \text{ for } \alpha_j \in W, \text{ and } \theta \geq 0 \tag{5.48}$$

Now, consider $\alpha_j \notin W$. Then, it is clear that

$$[v_0\ V^{\top}]\alpha_j > \check{c}_j. \tag{5.49}$$

Let $\mathbf{r}$ defined by

$$\mathbf{r}^T = [r_0 \; r_1 \; .. \; r_m] = \check{C}_D{}^T D^{-1}.$$

Then, if

$$\mathbf{r}^T \alpha_j < 0 \text{ for some } j,$$

we have

$$[v'_0 \; V'^T]\alpha_j = [v_0 \; V^T]\alpha_j + \theta \mathbf{r}^T \alpha_j \geq \check{c}_j \tag{5.50}$$

if and only if

$$\theta \leq -([v_0 \; V^T]\alpha_j - \check{c}_j)/\mathbf{r}^T\alpha_j.$$

From this, if

$$\mathbf{r}^T \alpha_j < 0 \text{ for some } j,$$

then

$$\theta \leq \theta_0 = \min \left\{ -([v_0 \; V^T]\alpha_j - \check{c}_j)/\mathbf{r}^T\alpha_j \,\middle|\, \mathbf{r}^T\alpha_j < 0 \right\}$$
$$< \infty \text{ (and } > 0).$$

If

$$\mathbf{r}^T \alpha_j \geq 0,$$

it is clear that $[v_0 \; V^T]^T$ will be a solution to the dual no matter how large θ becomes. In this case, we define

$$\theta_0 = \infty.$$

Hence, for $0 \leq \theta \leq \theta_0 \leq \infty$, we have

$$w' = [b_0 \; B^T][v'_0 \; V'^T]^T$$
$$= [v_0 \; V^T][b_0 \; B^T]^T + \theta \mathbf{r}^T [b_0 \; B^T]^T \text{ (from (5.50))}$$
$$= w + \theta \check{C}_D{}^T D^{-1}[b_0 \; B^T]^T = w + \theta z_a.$$

Since

$$z_a < 0, \text{ and } \theta > 0,$$

we have

$$w' < w.$$

This concludes the first part.

Next, from (5.40), for $j = 1,2,..,n$, we have

$$[v'_0 \ V'^{T}]\boldsymbol{\alpha}_j = [v_0 \ V^T]\boldsymbol{\alpha}_j + \theta_0 r^T \boldsymbol{\alpha}_j. \tag{5.51}$$

Note that for $j = 0,1,2,..,n$,

$$[v'_0 \ V'^{T}]\boldsymbol{\alpha}_j = v'_{sj} + \check{c}_j,$$

and (5.52)

$$[v_0 \ V^T]\boldsymbol{\alpha}_j = v_{sj} + \check{c}_j.$$

Hence, by substituting (5.52) into (5.51), we have

$$v'_{sj} = v_{sj} + \theta_0 \mathbf{r}^T \boldsymbol{\alpha}_j \text{ for } j = 0,1,..,n.$$

Therefore, by (5.47), we have

$$v'_{sj} = v_{sj} + \theta_0(\check{z}_j - \check{c}_j) \text{ for } j = 0,1,..,n.$$

Finally, the remainder of the proof can be immediately established. ■

From (5.42), there is at least one $v'_{sj} = 0$. This implies that at least one new vector $\boldsymbol{\alpha}_j$ in W can be used to enter the basis, and the process can be repeated until

$$\check{z}_j - \check{c}_j \geq 0 \text{ for all } \boldsymbol{\alpha}_j \text{ in } W.$$

Then, again, Theorem 5.10 can be used to iterate the same process until

$$\check{z}_j - \check{c}_j \geq 0 \text{ for all } j = 1,2,..,n.$$

In particular, we have the following algorithm.

DEFINITION 5.7. (**Primal-Dual Algorithm**) The following procedure is called the **Primal-Dual Algorithm**:

Initialization step: Start with a simplex tableau for the modified artificial problem (5.36) (some arrangement may be required.)

Main steps: Compute V_s by (5.37), and construct

$$W = \{\alpha_j | \ v_{sj} = 0\}.$$

1) If $\check{z}_j - \check{c}_j \geq 0$ for all $j = 1,2,..,n$, then stop; the original program is either optimal (if $z_a = 0$ and $v_0 = 0$) or unbounded (if $z_a < 0$ or $v_0 \neq 0$.)

Otherwise, if $\check{z}_j - \check{c}_j \geq 0$ for all $\alpha_j \in W$, then compute θ_0 by (5.41), V'_s by (5.42), update $V_s = V'_s$, and $W = \{\alpha_j | \ v_{sj} = 0\}$.

2) Drive a vector α_j in W with $\check{z}_j - \check{c}_j < 0$ into D; then drive out a basis vector, and update the current tableau both in the same way done in Phase I of the two-phase algorithm. Repeat Step 1.

The next example is given to illustrate the primal-dual algorithm.

EXAMPLE 5.11. Use the primal-dual algorithm to solve the following program:

$$\begin{array}{lrcl}
\text{Maximize} & -x_1 + 4x_2 + 2x_3 + 4x_4 & & \\
\text{Subject to} & x_1 + 4x_2 + x_3 + (1/2)x_4 & = & 10, \\
 & 2x_1 + 5x_2 + x_3 + x_4 & = & 14. \\
 & x_1, x_2, x_3, x_4 \geq 0. & &
\end{array}$$

SOLUTION. We first obtain the modified problem:

Maximize $\quad - x_1 + 4x_2 + 2x_3 + 4x_4$

$$\begin{aligned} \text{Subject to} \quad x_0 + x_1 + x_2 + x_3 + x_4 &= b_0, \\ x_1 + 4x_2 + x_3 + (1/2)x_4 &= 10, \\ 2x_1 + 5x_2 + x_3 + x_4 &= 14. \\ x_0, x_1, x_2, x_3, x_4 &\geq 0. \end{aligned}$$

By applying the primal-dual algorithm, we generate the the following tableaux:

TABLEAU I			0	-1	4	2	4
$\check{C}_D$	X_D	Basis	α_0	α_1	α_2	α_3	α_4
-1	b_0	$\mathbf{e}_1$	1	1	1	1	1
-1	10	$\mathbf{e}_2$	0	1	4	1	1/2
-1	14	$\mathbf{e}_3$	0	2	5	1	1
z_a	$-24-b_0$	$\check{z}_j-\check{c}_j$	-1	-4	-10	-3	-5/2
w	-	V_{sj}	4	5	0	2	0

TABLEAU II			0	-1	4	2	4
$\check{C}_D$	X_D	Basis	α_0	α_1	α_2	α_3	α_4
-1	b_0-5/2	$\mathbf{e}_1$	1	3/4	0	3/4	8/7
0	5/2	α_2	0	1/4	1	1/4	1/8
-1	3/2	$\mathbf{e}_3$	0	3/4	0	-1/4	3/8
z_a	1-b_0	$\check{z}_j-\check{c}_j$	-1	-3/2	0	-1/2	-5/4
w	-	v_{sj}	4	5	0	2	0

TABLEAU III			0	-1	4	2	4
$\check{C}_D$	X_D	Basis	α_0	α_1	α_2	α_3	α_4
-1	b_0-6	$\mathbf{e}_1$	1	-1	0	4/3	0
0	2	α_2	0	0	1	1/3	0
0	4	α_4	0	2	0	-2/3	1
z_a	6-b_0	$\check{z}_j-\check{c}_j$	-1	1	0	-4/3	0
w	-	v_{sj}	5/2	13/2	0	0	0

TABLEAU IV			0	-1	4	2	4
$\check{C}_D$	X_D	Basis	α_0	α_1	α_2	α_3	α_4
-1	b_0-14	$\mathbf{e}_1$	1	-1	-4	0	0
0	6	α_3	0	0	3	1	0
0	8	α_4	0	4	2	0	1
z_a	14-b_0	$\check{z}_j-\check{c}_j$	-1	1	4	0	0
w	-	v_{sj}	0	0	9	10	0

TABLEAU V			0	-1	4	2	4
$\check{C}_D$	X_D	Basis	α_0	α_1	α_2	α_3	α_4
0	b_0-14	α_0	1	-1	-4	0	0
0	6	α_3	0	0	3	1	0
0	8	α_4	0	4	2	0	1
z_a	0	$\check{z}_j-\check{c}_j$	0	0	0	0	0
w	44	v_{sj}	0	0	9	10	0

Graphically, the dual attains its minimal value 44 at the point $[-4\ 6]^T$. This result can be used to check with the maximal value of the primal: 2(6) + 4(8) = 44.

FINAL REMARK. 1. If a linear program is given an initial feasible solution, then it could be reasonable to compare the regular simplex algorithm with the preliminary primal-dual algorithm. Then, the larger the number of contraints are given, the more disadvantageous the latter algorithm would be. However, if this is the case, then we can use it to solve the dual (instead of the primal). The point is that the latter algorithm needs to drive out merely a number of artificial vectors that are relatively small, comparing with the number of variables of the original program.

2. If a linear program is not given an initial feasible solution, then it must be reasonable to compare the two-phase simplex algorithm with the primal-dual algorithm. Then, the latter algorithm is more advantegeous than the former one in the sense that it need only one phase. The disadvantage of the latter is that it has to deal with one more variable (from the additional constraint.) However, if the problem is given with a large number of variables (or constraints), then one in addition would not mean much.

Therefore, theoretically, and especially in regard of probability, the primal-dual algorithm would be superior to the previous ones.

BIBLIOGRAPHY

BIBLIOGRAPHY

1. Bazaraa, M. S., Sherali H. D., and Shetty C. M. Nonlinear Programming: Theory and Algorithm, John Wiley & Sons, Inc., New York, 1993.

2. Cooper, Leon and Steinberg, David, Methods and Applications of Linear Programming, W. B. Saunder Company, Philadelphia, 1993.

3. Gale, David, The Theory of Linear Economic Models, McGraw-Hill Book Company, Inc., New York, 1960.

4. Gass, Saul I., Linear Programming Methods and Applications, McGraw-Hill Book Company, Inc., New York, 1985.

5. Goldman, A. J., "Resolution and Separation Theorems for Polyhedral Convex Sets," Annals of Mathematical Studies 38 (1956), 41-51.

6. Goldman, A. J., and Tucker, A. W., "Polyhedral Convex Cones," Annals of Mathematical Studies 38 (1956), 19-40.

7. Goldman, A. J., and Tucker, A. W., "Theory of Linear Programming," Annals of Mathematical Studies 38 (1956), 53-97.

8. Graybill, Franklin A., Matrices with Applications in Statisics, Wadsworth Publishing Company, Inc., California, 1983.

9. Hadley G., Linear Programming, Addison-Wesley Publishing Company, Inc., Massachusetts, 1962.

10. Noble, Ben and Daniel, James W., Applied Linear Algebra, Prentice-Hall, New Jersey, 1988.

11. Spresser, Diane M., Convex Cones and Linear Inequalities, Master Thesis, University of Tennessee, 1967.

12. Tucker, A. W., "Dual Systems of Homogeneous Linear Relations," Annals of Mathematical Studies 38 (1956), 3-18.

VITA

The author was born in Angiang, Vietnam, on July 7, 1953. He was graduated from Thoai-Ngoc-Hau High School in 1972. Attending the University of Southwestern Louisiana, he studied Computer Science and Mathematics, and received a B. S. degree in Mathematics in 1986. In 1990, he entered the University of Tennessee, Knoxville to study for a master degree in Mathematics.