THE THEORY OF LINEAR PROGRAMMING

by

ELMER EARL BRANSTETTER

B. S., Central Missouri State College, 1961

A MASTER'S REPORT

submitted in partial fulfillment of the

requirements for the degree

MASTER OF SCIENCE

Department of Mathematics

KANSAS STATE UNIVERSITY
Manhattan, Kansas

1963

Approved by:

H. Thomas Parker
Major Professor

TABLE OF CONTENTS

Introduction

Programming may be defined as the construction of a schedule of actions by means of which an organization or complex of activities may move from a defined state toward some specifically defined objective. Such a schedule should explicitly prescribe the resources and the goods and services utilized, consumed, or produced in the accomplishment of the programmed actions. The resources and the goods and services utilized, consumed, or produced by the activities may be referred to as "commodities".

Linear Programming is the search for a program which will, in some sense, most nearly accomplish the desired objectives without exceeding stated resource limitations. To accomplish this, all interrelationships in the organization or complex of activities must be represented by a system of simultaneous linear relations in which the variables are the quantities of the activities to be performed, the coefficients are the requirements of each activity for each commodity, and each linear relation expresses the relationship between the sum of the requirements of all activities for a single commodity and the outputs of that commodity from all activities. In preparing such a program, it is necessary to insert in the system a specification of the initial status in terms of the quantities of each commodity on hand, any subsequent limitations, and a statement of the objectives. Thus, the determination of the desired program mathematically involves the solution of a system of linear relations, and hence, is called Linear Programming.

It is generally possible to determine the linear program using a mathematical model which will maximize the accomplishment of given objectives within those stated resource limitations. Alternately, it will be possible to determine the program which will minimize the requirements for any group of commodities needed to accomplish any fixed objective.

The origin and background of linear programming has been primarily in the field of econometrics and research into economic relationships. The history of linear programming dates from the publication of work of the methematical economist Leon Walras in 1874. Walras showed that the price of any number of commodities at a single time can be determined by solving simultaneously the correct number of equations in terms of the number of unknowns for which a solution is sought. At that time, the concept was revolutionary, but present-day methods of linear programming are completely different from those used by Walras. However, it was this first attempt to solve programming problems by stating the problem conditions in equation form that provides the connection between Walras and linear programming.

Linear programming, as it is known today, began with the input-output method of analysis developed by the economist Wassily W. Leontief in the 1920's. The present-day development stems primarily from the work of George B. Dantzig. Dantzig is credited with developing the Simplex Method of linear programming, which is essentially a method of solving simultaneous equations and inequalities for an optimum or best solution (a proof of the Simplex Method is given in the Appendix). Since Dantzig announced his development in 1947, the adaptation and application of the simplex method to solving linear programming problems has been fostered by simultaneous computational advances through high-speed digital computing machines.

It is to be emphasized that the simplex method is not the only linear-programming method. However, most known methods are derived or have evolved from the simplex method, which is considered to be the fundamental linear-programming method. Some of these methods are: the Index, the Modi, the Ratio Analysis, and the Symmetric Methods.

The aim of this report is to give a presentation of the theory of linear programming utilizing the concept of duality with respect to the maximization and minimization problems. A purely algebraic approach has been employed. Whenever possible, the more convenient matrix notation has been used to represent systems of linear relations. No attempt has been made to give a complete geometric interpretation of any result.

Dual Linear Programs

The following constitute a pair of <u>dual</u> <u>linear</u> programs (or dual linear programming problems):

I Maximize

(1.1) $$c_1x_1 + c_2x_2 + \cdots + c_nx_n$$

subject to the $m + n$ constraints,

(1.2) $$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &\leqq b_1 \\ &\ddots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &\leqq b_m \end{aligned}$$

(1.3) $$x_1 \geqq 0, \quad x_2 \geqq 0, \ldots, x_n \geqq 0.$$

II Minimize

(1.4) $$u_1b_1 + u_2b_2 + \cdots + n_mb_m$$

subject to the $n + m$ constraints,

(1.5) $$\begin{aligned} u_1a_{11} + u_2a_{21} + \cdots + u_ma_{m1} &\geqq c_1 \\ &\ddots \\ u_1a_{1n} + u_2a_{2n} + \cdots + u_ma_{mn} &\geqq c_n \end{aligned}$$

(1.6) $$u_1 \geqq 0, \quad u_2 \geqq 0, \ldots, u_m \geqq 0.$$

Here the a_{ij}, b_i, and c_j are given real numbers. The inequalities of (1.2) are called <u>row constraints</u> since they involve the rows of the matrix of a_{ij}'s; those of (1.5) are called <u>column constraints</u>. Problems of this type will later be referred to as problems in <u>canonical form</u>.

An interpretation of the maximization problem may be given in terms of Activity Analysis of Production. Let there be n activities, i.e., ways of making a single desired commodity from available stocks of m primary materials. Let a_{ij} be the amount of the i-th material used in one unit of the j-th activity, b_i the available stock of the i-th material, c_j the quantity of the desired commodity made by one unity of the j-th activity, and x_j the number of units of the j-th activity to be undertaken. The maximization problem is then a search for an activity vector $(x_1, x_2, \ldots, x_n)$ which will yield the greatest possible output (1.1) of the desired commodity, subject to the constraints (1.2) set by available stocks of the m materials and by the natural impossibility (1.3) of negative activity levels.

The dual problem pertains to accounting (fictitious or shadow) prices attached to the m materials, on a scale whose unit is the price of the desired commodity. In solving this problem, one seeks a price vector $(u_1, u_2, \ldots, u_m)$ that minimizes the total accounting value (1.4) of the available stocks of materials, subject to the requirement (1.5) that the accounting values of the quantity of the desired commodity made by one unit of an activity can never exceed the total accounting value of the materials used in that unit, and to the natural requirement (1.6) that all accounting prices be nonnegative.

Returning to the mathematical discussion, the dual linear programming problems will now be stated in the more compact matrix notation to be employed in the development of the theory in this report. All vectors will be from an n-dimensional real vector space v_n and will be denoted by upper case letters. The components of a vector will be denoted by lower case letters. The transpose of a matrix X will be denoted by X^T. In particular, let $U = (u_1, u_2, \ldots, u_m)$ and $C = (c_1, c_2, \ldots, c_n)$ be row vectors and let $X = (x_1, x_2, \ldots, x_n)$ and $B = (b_1, b_2, \ldots, b_m)$ be column vectors. However, vectors U, B, and C will be used interchangeably as row and column vectors. The use of inequality signs in terms of vectors will be as follows:

$X > Y$ means $x_i > y_i$ for all i,

$X \geqq Y$ means $x_i \geqq y_i$ for all i,

$X \geq Y$ means $x_i \geqq y_i$ for all i, $x_i > y_i$ for some i.

The basic problems for initial discussion then can be written in canonical form using matrix notation as:

(1.7) I Maximize CX, subject to $AX \leqq B$, $X \geqq 0$.

(1.8) II Minimize UB, subject to $UA \geqq C$, $U \geqq 0$.

A vector X satisfying the m + n constraints (1.2) and (1.3) without necessarily yielding the maximum in (1.1) will be called a <u>feasible</u> vector or solution for the maximization problem. A feasible vector X which provides the desired maximum for CX will be called an <u>optimal</u> vector or solution for the maximization problem.

The terms feasible solution and optimal solution are defined analogously for the minimization problem.

<u>Lemma 1.1.</u> If X and U are feasible, then $CX \leqq UB$.

Proof: Consider the constraints of the problems (1.7) and (1.8). Upon substituting and re-associating, it follows that $CX \leqq (UA)X = U(AX) \leqq UB$.

<u>Lemma 1.2.</u> If X^o and U^o are feasible, and $CX^o = U^oB$, then X^o and U^o are optimal.

Proof: If $CX^o = U^oB$, then by Lemma 1.1, $U^oB = CX^o \leqq UB$ for all feasible vectors U. Thus U^o is optimal. If $CX^o = U^oB$, by Lemma 1.1, $CX \leqq U^oB = CX^o$ for all feasible vectors X. Thus X^o is optimal.

According to Lemma 1.2, the assertion that X^o and U^o are a pair of optimal vectors can be checked directly in the schematic form of the dual programs:

1. One checks the feasibility of X^o by making sure that all its components are nonnegative and that the inner product of each row of the matrix A with X^o is not greater than the corresponding b_i.

2. One checks the feasibility of U^o by making sure that all its components are nonegative and that its inner product with each column of A^T with U^o is not less than the corresponding c_j.

3. Finally, one checks the equality of the inner product of C and X^o with that of U^o and B.

Lemma 1.3. If $B \geqq 0$, then the maximization problem has feasible constraints. Let the n columns of the matrix A be denoted by $A_1, \ldots, A_n$. Then, if $A_i \geqq 0$ for each i, the minimization problem has feasible constraints.

Proof: The validity of the first statement is seen by noting that the zero vector is feasible for the maximization problem. The second statement is proved by noting that under the hypothesis, any vector U will be feasible if all its components are sufficiently large.

The following two examples illustrate the possibility that constraints in one or both of the dual problems may not be feasible.

Example 1.

$$\begin{aligned} x_1 - 3x_2 &\leqq 2 \\ 4x_1 - 2x_2 &\leqq 0 \\ x_1 \geqq 0, \quad x_2 &\geqq 0. \end{aligned} \qquad\qquad \begin{aligned} u_1 + 4u_2 &\geqq 3 \\ -3u_1 - 2u_2 &\geqq 1 \\ u_1 \geqq 0, \quad u_2 &\geqq 0. \end{aligned}$$

Here, $X = (1,2)^T$ is feasible but the column constraints clearly cannot be met, so the minimization has no feasible solution vector.

Example 2.

$$\begin{aligned} 0x_1 + x_2 &\leqq -1 \\ 0x_1 + 4x_2 &\leqq 3 \\ x_1 \geqq 0, \quad x_2 &\geqq 0. \end{aligned} \qquad\qquad \begin{aligned} 0u_1 + 0u_2 &\geqq 1 \\ u_1 + 4u_2 &\geqq 2 \\ u_1 \geqq 0, \quad u_2 &\geqq 0. \end{aligned}$$

Here, neither the row constraints or the column constraints can be satisfied, so neither problem has a feasible solution vector.

Dual Systems of Homogeneous Linear Relations

In this section, theorems concerning dual systems of homogeneous linear relations are derived. These theorems play an important role in the development of duality and existence theorems of the next section. Throughout this section, the letters A, B, C, and D will denote m x n matrices with arbitrary real elements. Also, the letters U and V will denote m-dimensional vectors from the vector space V_n while the letters X and Y will denote n-dimensional vectors from V_n.

First, consider the pair of dual systems $A^TU \geqq 0$ and $AX = 0$, $X \geqq 0$ where $A = (A_1, A_2, \ldots, A_n)_{m \times n}$ is an n-columned matrix with arbitrary real elements.

Lemma 2.1. The dual systems $A^TU \geqq 0$ and $AX = 0$, $X \geqq 0$ possess solutions U and X such that $A_i^TU + x_i > 0$, where A_i is any one of the n columns of A.

Proof: The proof is by induction on n, the number of columns in A. The initial case, n = 1, is trivial: If $A_i = 0$, take $U = 0$ and $x_i = 1$: if $A_i \neq 0$, take $U = A_i$ and $x_i = 0$.

Now assuming that the Lemma holds for a matrix A of n columns, it will be shown that the Lemma holds for a matrix $\bar{A} = (A, A_{n+1}) = (A_1, \ldots, A_n, A_{n+1})$ of n + 1 columns. Applying the Lemma to A, one obtains vectors U and X such that $A^TU \geqq 0$, $AX = 0$, $X \geqq 0$, and $A_i^TU + x_i > 0$.

If $A_{n+1}^TU \geqq 0$, take $\bar{X} = (X, 0)^T$. Then $\bar{A}^TU \geqq 0$, $\bar{A}\bar{X} = 0$, $\bar{X} \geqq 0$, $A_i^TU + x_i > 0$, which extends the Lemma to the matrix $\bar{A}$.

However, if $A_{n+1}^T U < 0$, an application of Lemma 2.1 to a particular $n \times m$ matrix $B = (B_1, B_2, \ldots, B_n)$ will again extend the Lemma to the matrix A.

Let $B = (B_1, \ldots, B_n) = (A_1 + k_1 A_{n+1}, \ldots, A_n + k_n A_{n+1})$, where the constants k_j are given by $k_j = -A_j^T U / A_{n+1}^T U$ for $j = 1, \ldots, n$. Note that the k_j are nonnegative since $A^T U \geqq 0$. This application of the Lemma to the matrix B yields vectors V and Y such that $B^T V \geqq 0$, $BY = 0$, $Y \geqq 0$, and $B_1^T V + y_1 > 0$. Then, if $\bar{Y}$ is chosen to be $\bar{Y} = (Y, \sum_{j=1}^{n} k_j y_j)^T$, it follows $\bar{Y} \geqq 0$, since $Y \geqq 0$ and $\sum_{j=1}^{n} k_j y_j \geqq 0$. Furthermore, $\bar{A}\bar{Y} = BY = 0$. This can be seen by writing

$$\bar{A}\bar{Y} = \left(\sum_{j=1}^{n} a_{1j} y_j + a_{1,n+1} \sum_{j=1}^{n} k_j y_j, \ldots, \sum_{j=1}^{n} a_{nj} y_j + a_{m,n+1} \sum_{j=1}^{n} k_j y_j \right)^T$$

$$= \left(\sum_{j=1}^{n} (a_{1j} + k_j a_{1,n+1}) y_j, \ldots, \sum_{j=1}^{n} (a_{mj} + k_j a_{m,n+1}) y_j \right)^T$$

$$= BY,$$

and by noting that $BY = 0$ by hypothesis.

Now, let $W = V + rU$, so that if r is chosen to be $-A_{n+1}^T V / A_{n+1} U$, then $\bar{A}^T W \geqq 0$. The latter relation follows by writing

$$A^T W = (A_1^T W, \ldots, A_n^T W)^T$$

$$= \left(A_1^T V - \frac{A_{n+1}^T V}{A_{n+1}^T U} A_1^T U, \ldots, A_n^T V - \frac{A_{n+1}^T V}{A_{n+1}^T U} A_n^T U \right)^T$$

$$= (A_1^T V - \frac{A_1^T U}{A_{n+1}^T U} A_{n+1}^T V, \ldots, A_n^T V - \frac{A_n^T U}{A_{n+1}^T U} A_{n+1}^T V)^T$$

$$= (A_1^T V + k_1 A_{n+1}^T V, \ldots, A_n^T V + k_n A_{n+1}^T V)^T$$

$$= B^T V \geqq 0,$$

and by noting that

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

Therefore, $A_i^T W + y_i = B_i^T V + y_i > 0$ for $i = 1, \ldots, n$.

Thus, by means of the vectors W and Y, the Lemma has been extended to the matrix $\bar{A} = (A, A_{n+1})$. Therefore, the induction on n is fully established and the Lemma must hold for all n.

The following corollary is a statement of a theorem first proved in 1902 by J. Farkas (9).

<u>Corollary 2.1.</u> If the inequality $A_o^T U \geqq 0$ holds for all solutions U of the system $A^T U \geqq 0$, then $A_o = AX$ for some $X \geqq 0$.

Proof: Consider the matrix $(-A_o, A)^T$ whose first column is $-A_o$, and the dual systems $(-A_o, A)^T U \geqq 0$ and $(-A_o, A)\ (x_o, X)^T = 0$, $(x_o, X)^T \geqq 0$. Note that $-A_o$ and x_o assume the roles of A_i and x_i with $i = 1$ in Lemma 2.1.

By this Lemma, one is assured of some solution U and $(x_o, X)^T$ such that $-A_o^T U \geqq 0$, $A^T U \geqq 0$, $-A_o x_o + AX = 0$, $x_o \geqq 0$, $X_o \geqq 0$, and $-A_o^T U + x_o > 0$. The relation $-A_o^T U + x_o > 0$ implies that $x_o > 0$, since by hypothesis of the Corollary, $A_o^T U \geqq 0$ for all solutions U of the system. Thus, from the relations $-A_o x_o + AX = 0$ and $x_o > 0$, it follows that $A_o = AX^o$, where $X^o = X/x_o \geqq 0$.

It is to be emphasized that the vector A_o in the above corollary is actually a nonnegative linear combination of the vectors $A_1, \ldots, A_n$, that is,

$$A_o = A_1 x_1/x_o + A_2 x_2/x_o + \cdots + A_n x_n/x_o,$$

$$= A_1 y_1 + A_2 y_2 + \cdots + A_n y_n,$$

with $y_i \geqq 0$, for $i = 1, \ldots, n$.

With this remark, the corollary may be interpreted as follows: if a linear inequality is dependent upon a system of linear inequalities, in the sense that it is satisfied by all solutions of the system, then the coefficients of that inequality depend linearly on the coefficients of the system, and furthermore, in this dependence only nonnegative coefficients need be used.

The following theorem is a key result because from it virtually all of the results in duality theory follow.

<u>Theorem 2.1.</u> The dual systems $A^T U \geqq 0$ and $AX = 0$, $X \geqq 0$

possess solutions U^* and X^* such that $A^TU^* + X^* > 0$.

Proof: By Lemma 2.1, there exist pairs of solutions U^j and X^j such that $A^TU^j \geqq 0$, $AX^j = 0$, $X^j \geqq 0$, and $A_j^TU^j + x_j^j > 0$. Let $U^* = \sum_{j=1}^{n} U^j$ and $X^* = \sum_{j=1}^{n} X^j$. Then, $A^TU^* = \sum_j A^TU^j \geqq 0$, $AX^* = \sum_j AX^j = 0$, and $X^* = \sum_j X^j \geqq 0$. Also, for $j = 1,\ldots,n$,

$$A_j^TU^* + x_j^* = \sum_{k=1}^{n} (A_j^TU^k + x_j^k) \geq A_j^TU^j + x_j^j > 0,$$

since $A^TU^k \geqq 0$ and $X^k \geqq 0$ for $k = 1,\ldots,n$. Therefore, $A^TU^* + X^* > 0$.

It is interesting to note that the solutions U^* and X^* of Theorem 2.1 have the property that for $j = 1,\ldots,n$, the j-th component of one of the vectors X^* and A^TU^* is zero if, and only if, the j-th component of the other is positive.

Corollary 2.2. The system of equations $AX = 0$ has (i) a solution $X > 0$ if there is no solution U such that $A^TU \geq 0$, and (ii) a solution $X \geq 0$ if there is no solution U such that $A^TU > 0$.

Proof: By Theorem 2.1, there exist solutions U^* and X^* such that $A^TU^* \geqq 0$, $AX^* = 0$, $X^* \geqq 0$, and $A^TU^* + X^* > 0$. From the hypothesis of Theorem 2.1, every solution U must satisfy the relation $A^TU \geqq 0$. Thus, if it is not true that $A^TU \geq 0$ for every solution U, then the mutually exclusive alternative $A^TU = 0$ must be true. In particular, $A^TU^* = 0$ if $A^TU^* \geq 0$ does not hold. Since U^* also satisfies

$A^TU^* + X^* > 0$, it follows that $X^* > 0$. This establishes part (i). Now, if any solution X of the system $AX = 0$, $X \geqq 0$ does not satisfy the relation $X \gtrsim 0$, then X must be zero. For the particular solution X^*, this implies that $A^TU^* > 0$, since $A^TU^* + X^* > 0$. The contrapositive of this result establishes part (ii), and hence, completes the proof.

<u>Theorem 2.2.</u> The dual systems $A^TU \geqq 0$, $B^TU = 0$ and $AX + BY = 0$, $X \geqq 0$ possess solutions U^* and X^*, Y^* such that $A^TU^* + X^* > 0$.

Proof: The application of the results of Theorem 2.1 to the dual systems $(A,B,-B)^TU \geqq 0$ and $(A,B,-B)(X,Y_1,Y_2)^T = 0$, $(X,Y_1,Y_2)^T \geqq 0$ implies the existence of solutions U^* and X^*, Y_1^*, Y_2^* such that

$$A^TU^* \geqq 0, \quad B^TU^* \geqq 0, \quad -B^TU^* \geqq 0, \quad AX^* + BY_1^* - BY_2^* = 0,$$

and

$$X^* \geqq 0, \quad Y_1^* \geqq 0, \quad Y_2^* \geqq 0,$$

and

$$A^TU^* + X^* > 0, \quad B^TU^* + Y_1^* > 0, \quad -B^TU^* + Y_2^* > 0.$$

If one chooses $Y^* = Y_1^* - Y_2^*$, then, clearly,

$$A^TU^* \geqq 0, \quad B^TU^* = 0, \quad AX^* + BY^* = 0, \quad X^* \geqq 0,$$

and $A^TU^* + X^* > 0$.

<u>Corollary 2.3.</u> Let the dual systems

$$A^TU \geqq 0, \quad B^TU = 0 \quad \text{and} \quad AX + BY = 0, \quad X \geqq 0$$

have the partitioned representation

$A^{1T} \geq 0$, $A^{2T} \geqq 0$, $B^T U = 0$ and $A^1 X_1 + A^1 X_2 + BY = 0$, $X_1 \geqq 0$, $X_2 \geqq 0$, where A^1 is any nonvacuous set of columns of A and A^2 is the set (possibly vacuous) of remaining columns of A. Then (i) <u>either</u> the left system has a solution U such that $A^{1T}U \geq 0$ <u>or</u> the right system has a solution X such that $X_1 > 0$. Also, (ii) <u>either</u> the left system has a solution U such that $A^{1T}U > 0$ <u>or</u> the right system has a solution X such that $X_1 \geq 0$.

Proof: Theorem 2.2 asserts the existence of solutions U^* and X_1^*, X_2^*, Y^* such that

$$A^{1T}U^* \geqq 0, \quad A^{2T}U^* \geqq 0, \quad B^T U^* = 0, \quad A^1 X_1^* + A^2 X_2^* + BY = 0,$$

and

$$X_1^* \geqq 0, \quad X_2^* \geqq 0, \quad A^{1T}U^* + X_1^* > 0, \quad A^{2T}U^* + X_2^* > 0.$$

For any solution U to satisfy the relation $A^{1T}U \geqq 0$, one of the mutually exclusive alternatives, $A^{1T} = 0$ or $A^{1T} \geq 0$, must hold. If for the particular solution U^*, $A^{1T}U^* = 0$, then $X_1^* > 0$, since $A^{1T}U^* + X_1^* > 0$. This establishes part (i). Also, any solution X must satisfy the relation $X \geqq 0$, and hence, must satisfy either $X \geq 0$ or $X = 0$. It follows that either $X_1 \geq 0$ or $X_1 = 0$, and, in particular, $X_1^* \geq 0$ or $X_1^* = 0$ must hold. If $X_1^* = 0$, then $A^{1T}U^* > 0$, since $A^{1T}U^* + X_1^* > 0$. This establishes part (ii).

<u>Theorem 2.3.</u> The dual systems $V \geqq 0$, $C^T V \geqq 0$ and $-CX \geqq 0$, $X \geqq 0$ possess solutions V^* and X^* such that

$V^* - CX^* > 0$ and $C^TV^* + X^* > 0$.

Proof: Let I denote the $m \times n$ identity matrix and apply the results of Theorem 2.1 to the dual systems

$$(I,C)^T V \geqq 0 \quad \text{and} \quad (I,C)(W,X)^T = 0, \quad (W,X)^T = 0,$$

where W is an $m \times 1$ column vector. This application implies the existence of solutions X^*, W^*, and V^* such that $(I,C)^T V^* \geqq 0$, $(I,C)(W^*,X^*)^T = 0$, $(W^*,X^*)^T \geqq 0$, and $(I,C)^T V^* + (W^*,X^*)^T > 0$. These relations when simplified, yield

(2.9) $V^* \geqq 0$, $C^TV^* \geqq 0$, $-CX^* = W^* \geqq 0$, $X^* \geqq 0$, $V^* - CX^* > 0$, and $C^TV^* + X^* > 0$.

Corollary 2.4. The dual system

$V \geqq 0$, $C^TV \geqq 0$ and $-CX \geqq 0$, $X \geqq 0$

possess solutions V^* and X^* for which the following alternatives hold:

(i) either $C^TV^* \geq 0$ or $X^* > 0$,

(ii) either $C^TV^* > 0$ or $X^* \geq 0$,

(iii) either $V^* > 0$ or $-CX^* \geq 0$,

(iv) either $V^* \geq 0$ or $-CX^* > 0$.

Proof: Statements (i) through (iv) are immediate consequences of relations in statement (2.9). It should be noted, however, that from the first and third relations of (2.9), it follows that $0 \leqq (C^TV)^T X = V^T(CX) \leqq 0$, and hence, $V^TCX = 0$ for all solutions V^* and X^*. Therefore, the alternatives (i) through (iv) are mutually exclusive.

<u>Theorem 2.4.</u> The general dual systems

$$
\begin{array}{ll}
(U \text{ unrestricted}) & -AX - BY = 0 \\
V \geqq 0 & -CX - DY \geqq 0 \\
A^TU + C^TV \geqq 0 & X \geqq 0 \\
B^TU + D^TV = 0 & (Y \text{ unrestricted})
\end{array}
$$

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

$$V^* - CX^* - DY^* > 0 \quad \text{and} \quad A^TU + C^TV^* + X^* > 0.$$

Proof: Applying the results of Theorem 2.3 to the dual systems

$$
\begin{bmatrix} U_1 \\ U_2 \\ V \end{bmatrix} \geqq 0, \quad
\begin{bmatrix} -A & -B & B \\ A & B & -B \\ C & D & -D \end{bmatrix}^T
\begin{bmatrix} U_1 \\ U_2 \\ V \end{bmatrix} \geqq 0 \quad \text{and} \quad
-\begin{bmatrix} -A & -B & B \\ A & B & -B \\ C & D & -D \end{bmatrix}
\begin{bmatrix} X \\ Y_1 \\ Y_2 \end{bmatrix} \geqq 0, \quad
\begin{bmatrix} X \\ Y_1 \\ Y_2 \end{bmatrix} \geqq 0,
$$

where U_1 and U_2 are $m \times 1$ column vectors, and where Y_1 and Y_2 are $n \times 1$ column vectors, implies that there exist solutions

$$U_1^* \geqq 0, \quad U_2^* \geqq 0, \quad V^* \geqq 0 \quad \text{and} \quad X^* \geqq 0, \quad Y_1^* \geqq 0, \quad Y_2^* \geqq 0,$$

such that

$$
\begin{array}{ll}
-A^TU_1^* + A^TU_2^* + C^TV^* \geqq 0 & AX^* + BY_1^* - BY_2^* \geqq 0 \\
-B^TU_1^* + B^TU_2^* + D^TV^* \geqq 0 & -AX^* - BY_1^* + BY_2^* \geqq 0 \\
B^TU_1^* - B^TU_2^* - D^TV^* \geqq 0 & -CX^* - DY_1^* + DY_2^* \geqq 0 \\
V^* - CX^* - DY_1^* + DY_2^* > 0 & -A^TU_1^* + A^TU_2^* + C^TV^* + X^* > 0.
\end{array}
$$

Choose $U^* = U_1^* - U_2^*$ and $Y^* = Y_1^* - Y_2^*$. Then,

$$A^TU^* + C^TV^* \geqq 0 \qquad -AX^* - BY^* = 0$$

$$B^TU^* + D^TV^* = 0 \qquad -CX^* - DY^* \geqq 0$$

$$A^TU^* + C^TV^* + X^* > 0 \qquad V^* - CX^* - DY^* > 0$$

Therefore, the vectors U^* and Y^* are the desired solutions.

Let the matrix C of Theorem 2.3 be a skew-symmetric matrix, i.e., $C^T = -C$. Then, the dual systems of Theorem 2.3 become $V \geqq 0$, $CV \leqq 0$ and $X \geqq 0$, $CX \leqq 0$ with solutions V^* and X^* such that $V^* - CX^* > 0$ and $X^* - CV^* > 0$. Since these two dual systems are the same, the systems corresponding to a skew-symmetric matrix is called <u>self-dual</u>.

<u>Theorem 2.5.</u> The self-dual system $KW \geqq 0$, $W \geqq 0$ where $K^T = -K$, has a solution W^* such that $KW^* + W^* > 0$.

Proof: The proof involves an application of the results of Theorem 2.3 to the skew-symmetric matrix $C = K^T = -K$. According to Theorem 2.3, there exist solutions V^* and W^* such that

$$V^* \geqq 0, \quad KV^* \geqq 0, \quad KX^* \geqq 0, \quad X^* \geqq 0, \quad V^* + KX^* > 0, \quad KV^* + X^* > 0.$$

Hence,

$$K(V^* + X^*) \geqq 0, \quad (V^* + X^*) \geqq 0, \text{ and } K(V^* + X^*) + (V^* + X^*) > 0.$$

If W^* is taken to be $V^* + X^*$, the Theorem follows.

More specifically, the two vectors W^* and KW^* have the property that for each $j = 1,\ldots,n$, the j-th component of one vector is positive if, and only if, the j-th component of the other is zero. Indeed, Theorem 2.5 shows their sum is positive, but their inner product $(W^*)^T(KW^*)$ is zero, since the transpose of $(W^*)^TKW^*$ is its own negative.

Definition. A slack inequality in a system of inequalities is an inequality $(\geqq 0)$ which is satisfied as a strict inequality (>0) by some solution of the system.

Now, consider the system $A^TU \geqq 0$, $B^TU = 0$ of Theorem 2.2. Let U be a solution of this system. It may happen that $A_i^TU > 0$ for some $i = 1,\ldots,m$. That is, $A_i^TU \geqq 0$ may be a slack inequality for some row A_i^T. If there does exist a solution U such that $A_i^TU \geqq 0$ is satisfied as a strict inequality, let this solution be denoted by U_i, i.e., $A_i^TU_i > 0$. Now, let I denote the set of indices i such that $A_i^TU_i > 0$ for some solution U of the system. Then $\sum U_i$, summed for all i in I, is a solution of the system. Indeed, $A^T(\sum U_i) \geqq 0$ and $B^T(\sum U_i) = 0$ since each U_i for i in I is a solution of the system. Furthermore, $A_i^T(\sum U_i) > 0$ for all i in I since at least one of the summands is positive. This shows that the slack inequalities in a system can be characterized collectively as the maximum set of inequalities of the system which are satisfied as strick inequalities by some solution of the system. Hence, the remaining inequalities in the system are those which are satisfied as equations by all solutions of the system.

Theorem 2.6. In the general dual systems

$$(2.10)\qquad \begin{array}{ll} (U \text{ unrestricted}) & -AX - BY = 0 \\ V \geqq 0 & -CX - DY \geqq 0 \\ A^TU + C^TV \geqq 0 & X \geqq 0 \\ B^TU + D^TV = 0 & (Y \text{ unrestricted}), \end{array}$$

of Theorem 2.4, each of the $m + n$ pairs of corresponding inequalities

$$(2.11)\qquad \begin{array}{lll} v_i \geqq 0 & -\sum_{j=1}^{n} c_{ij}x_j - \sum_{k=1}^{m} d_{ik}y_k \geqq 0, & (i = 1,\dots,m) \\ \sum_{h=1}^{m} a_{hj}u_h + \sum_{i=1}^{m} c_{ij}v_i \geqq 0 & x_j \geqq 0 & (j = 1,\dots,n) \end{array}$$

contains exactly one inequality that is slack relative to the system.

Proof: Let U, V and X, Y be any solution of the given dual system. Then, by multiplying together corresponding terms in the dual systems and summing, the relations

$$(2.12)\qquad U^T(-AX - BY) = -U^TAX - U^TBY = 0,$$

$$(2.13)\qquad V^T(-CX - DY) = -V^TCX - V^TDY \geqq 0,$$

$$(2.14)\qquad (A^TU + C^TV)^TX = U^TAX + V^TCX \geqq 0,$$

$$(2.15)\qquad (B^TU + D^TV)^TY = U^TBY + V^TDY = 0,$$

are obtained. Adding (2.12) to (2.14) and (2.13) to (2.15), one obtains $-U^TBY + V^TCX \geqq 0$ and $U^TBY - V^TCX \geqq 0$. Hence, $U^TBY - V^TCX = 0$. Using this result, equations (2.12) and (2.15) yield the equations,

$$(2.16)\qquad -U^TAX = U^TBY = V^TCX = -V^TDY.$$

substituting from (2.16) into (2.13) and (2.14), it is evident that

$$(2.17)\qquad \begin{aligned} V^T(-CX - DY) &= \sum_i v_i\Big(-\sum_j c_{ij}x_j - \sum_k d_{ik}y_k\Big) = 0, \\ (A^TU + C^TV)^TX &= \sum_j \Big(\sum_h a_{hj}u_h + \sum_i c_{ij}v_i\Big)x_j = 0. \end{aligned}$$

Equation (2.17) shows that in each pair of corresponding inequalities of (2.11), at least one sign of equality must hold for all solutions of the systems; otherwise, $V(-CX - DY) \geq 0$ or $(A^T + C^TV)^TX \geq 0$ for some solutions. Therefore, each pair of dual inequalities in (2.11) contains at most one inequality that is slack.

By Theorem 2.4, there exists solutions U^*, V^* and X^*, Y^* of the dual systems such that

$$\begin{aligned} v_i^* + \Big(-\sum_j c_{ij}x_j^* - \sum_k d_{ik}y_k^*\Big) &> 0 \qquad (i=1,\ldots,m), \\ \Big(\sum_h a_{hj}u_h^* + \sum_i c_{ij}v_i^*\Big) + x_j^* &> 0 \qquad (j=1,\ldots,n). \end{aligned}$$

Hence, each pair of corresponding inequalities contains at least one inequality that is slack.

Taken together, the last two paragraphs imply that each pair of corresponding inequalities in the general dual systems (2.10) contains exactly one inequality that is slack relative to the system.

The property of the system (2.10) exhibited in Theorem 2.6 can be described collectively as complementary slackness, i.e., the set of slack inequalities in the one system is exactly complementary to the set of slack inequalities in the other system. The property of complementary slackness applies, of course, to the pairs of dual systems in Theorems 2.1, 2.2, and 2.3. Indeed, Theorem 2.6 can be reduced to these pairs of dual systems by letting certain of the matrices A, B, C, and D become vacuous. In particular, this property shows that the alternatives in the various parts of Corollaries 2.2, 2.3, and 2.4 are mutually exclusive.

Theorem 2.7. In the self-dual system $KW \geqq 0$, $W \geqq 0$, $K^T = -K$; each of the n pairs of corresponding inequalities $\sum_j k_{ij} w_j \geqq 0$ and $w_i \geqq 0$, $(i = 1, \ldots, n)$, contains exactly one inequality that is slack relative to the self-dual system.

Proof: For any solution W of the given system, $W^T K W = 0$, because $W^T K W = -W^T K W$. Hence, in each of the pairs of corresponding inequalities $\sum_j k_{ij} w_j \geqq 0$ and $w_i \geqq 0$, $(i = 1, \ldots, n)$, at least one sign of equality must hold for all solutions W. Therefore, each pair of corresponding inequalities contains at most one inequality that is slack.

By Theorem 2.5, there exists a solution W^* such that

$$\sum_j k_{ij} w_j^* + w_j^* > 0, \quad (i = 1, \ldots, n).$$

Hence, each pair of corresponding inequalities of the self-dual system contains exactly one inequality that is slack relative to the self-dual system.

It should be emphasized that a system of inequalities, $AX \leqq B$, in which are included the inequalities, $X_1 \geqq 0, \ldots, X_n \geqq 0$, restricting the unknown of a solution to nonnegative values is necessarily nonsingular, since no linear dependence of columns of the matrix A can occur. Therefore, linear programming constraints, as normally formulated, are nonsingular systems. However, the solution of the basic system of inequalities may be viewed in two ways: (1) to solve a system $(A,I)X \leqq B$ of $m + n$ linear relations in n unknown, or alternately, (2) to find nonnegative solutions of a system, $AX \leqq B$, $X \geqq 0$, of m linear relations in n unknowns. The problem has been studies in this section from the latter viewpoint.

Duality and Existence Theorems

Two of the fundamental results in linear programming theory are the Duality Theorem and the Existence Theorem. The essential step in the derivation of these theorems is the use of Theroem 2.5 to prove two lemmas that lead at once to the basic theorems about the dual pair of linear programming problems (1.7) and (1.8):

(3.1) Maximize CX, subject to $AX \leqq B$, $X \geqq 0$.

(3.2) Minimize UB, subject to $UA \geqq C$, $U \geqq 0$.

The first step is to use matrices A, B, and C to construct a skew-symmetric matrix of order $m + n + 1$:

$$K = \begin{bmatrix} 0 & -A & B \\ A^T & 0 & -C^T \\ -B^T & C & 0 \end{bmatrix} = -K^T.$$

By Theorem 2.5, there is an $(m + n + 1) \times 1$ column vector $W = (U_o, X_o, t_o)^T \geqq 0$, where t_o is some scalar, such that

(3.3) $W \geqq 0$, $KW \geqq 0$, and $KW + W > 0$.

When the expressions for K and W are substituted in (3.3), the following six inequalities in nonnegative variables are obtained:

$$(3.4) \qquad \begin{aligned} &\text{(i)} && AX_o \leqq Bt_o \\ &\text{(ii)} && U_oA \geqq t_oC \\ &\text{(iii)} && CX_o \geqq BU_o \\ &\text{(iv)} && AX_o < t_oB + U_o^T \\ &\text{(v)} && t_oC^T < A^TU_o^T + X_o \\ &\text{(vi)} && B^TU_o^T < CX_o + t_o. \end{aligned}$$

Throughout this section, the symbols W, U_o, X_o, t_o, and inequalities (i) through (vi) will have the meanings just designated. Two alternatives occur in (3.4): either $t_o > 0$ or $t_o = 0$. Whether or not the systems (3.1) and (3.2) have optimal solutions depends on whether $t_o > 0$ or $t_o = 0$. The effects of these two possibilities will be investigated separately in Lemmas 3.1 and 3.2.

Lemma 3.1. If $t_o > 0$, then the dual programs (3.1) and (3.2) have optimal vectors X^o and U^o such that $CX^o = U^oB$, $(U^o)^T + B > AX^o$, and $U^oA + (X^o)^T > C$.

Proof: Since $t_o > 0$, the nonnegative vector $W = (U_o, X_o^T, t_o)^T$ can be normalized so that $t_o = 1$ without affecting the validity of relations (3.4). Then, (i) and (ii) with $t_o = 1$ show that X_o/t_o and U_o/t_o are feasible vectors for their respective programs, while (iii) and Lemma 1.1 show that $C(X_o/t_o) = (U_o/t_o)B$. Hence, Lemma 1.2 implies that X_o/t_o and U_o/t_o are optimal solutions. Thus, one can choose X_o/t_o and U_o/t_o to be the desired vectors X^o and U^o. The strict inequalities of the Lemma follow respectively from (iv) and (v) with $t_o = 1$.

Lemma 3.2. Suppose $t_o = 0$. Then at least one of the dual programs has no feasible vector. If the maximization problem has a feasible vector, then the set F of all feasible vectors is unbounded and the function CX is not bounded above on F. Dually, if the minimization problem has a feasible vector, then F is unbounded and UB is not bounded below on F. Neither of the dual problems has an optimal vector.

Proof: Assume that X and U are feasible vectors for the maximization and minimization problems respectively. With $t_o = 0$, inequalities (i), (ii), and (vi) read $AX_o \leqq 0$, $0 \leqq U_oA$, and $CX_o > U_oB$. It follows from (ii), (vi), the fact that $U_o \geqq 0$, and the maximization problem constraints $AX \leqq B$, that

$$(3.5) \qquad 0 \leqq U_oAX \leqq U_oB < CX_o.$$

Now, with the aid of (i), the fact that $X_o \geqq 0$, and the minimization problem constraints $UA \geqq C$, it follows that $0 \geqq UAX_o \geqq CX_o$. This clearly contradicts (3.5). So, $t_o = 0$ precludes the possibility that both of the dual programs are feasible. Sometimes, neither program is feasible.

Now, the situation where one, and hence, only one, of the dual programs is feasible must be studied. Suppose, for example, that X is a feasible vector for the maximization problem. To prove the second statement of the Lemma, one must examine the infinite ray consisting of the vectors $X + rX_o$, where r is a positive scalar. It is a straightforward calculation to show that $X + rX_o$ are feasible. Indeed, $X + rX_o \geqq 0$, and using (i) with $t_o = 0$, it is evident that $A(X + rX_o) \leqq AX \leqq B$. Thus, the entire infinite ray consists of feasible vectors. Furthermore, the fact that $CX_o > 0$, as was shown in (3.5), makes it clear that $C(X + rX_o) = CX + rCX_o$ can be made arbitrarily large by choosing r large enough. Thus, the maximization problem with a feasible solution has an infinite number of feasible solutions, but has no maximum when $t_o = 0$. An analogous discussion shows that if the minimization problem is the only feasible one, then the desired minimum is not attainable.

In summary, if $t_o = 0$, then at least one of the dual programs is not feasible and neither has an optimal solution.

Corollary 3.1. Either both the maximization and minimization problems have optimal vectors or neither does. In the first case, the attained maximum and minimum are equal; their common magnitude is called the optimal value of the dual programs.

Proof: If one of the dual programs has an optimal vector, then the last statement of Lemma 3.2 implies that $t_o > 0$, and then Lemma 3.1 implies that both programs have optimal vectors X^o and U^o such that the maximum CX^o is equal to the minimum U^oB.

Corollary 3.2. The dual programs have an optimal value if, and only if, either CX or UB is bounded on the corresponding nonvacuous set of feasible vectors.

Proof: the necessity of the condition is clear. To prove the sufficiency, suppose that the maximization problem has a non-vacuous set F of feasible vectors and that CX is bounded on F.

The second statement of Lemma 3.2 implies that $t_0 > 0$. Thus, by Lemma 3.1, both of the dual programs have optimal vectors, and according to Corollary 3.1, they have an optimal value.

Theorem 3.1 (Duality Theorem). A feasible vector X^o for the maximization problem is optimal if, and only if, there is a feasible vector U^o for the minimization problem with $U^oB = CX^o$. Alternately, a feasible vector U^o for the minimization problem is optimal if, and only if, there is a feasible vector X^o for the minimization problem with $CX^o = U^oB$.

Proof: To prove the necessity of the first stated condition, suppose that X^o is an optimal vector for the maximization problem. Then, the last statement of Lemma 3.2 implies that $t_0 > 0$. Thus, by Lemma 3.1, the minimization problem also has an optimal vector. According to Corollary 3.1, the attained maximum, CX^o and the minimum U^oB are equal. That the first stated condition is also sufficient follows immediately from Lemma 3.2. Proof of the second statement of the Theorem is similar.

Theorem 3.2 (Existence Theorem). A necessary and sufficient condition that one, and hence both, of the dual problems have optimal vectors is that each problem has a feasible vector.

Proof: The necessity of the stated condition is evident from the Duality Theorem.

If each of the dual problems has a feasible vector, it follows from the first statement of Lemma 3.2 that $t_o > 0$. If $t_o > 0$, Lemma 3.1 implies the existence of optimal vectors X^o and U^o. This proves the sufficiency of the condition.

The following corollary deals with the way in which a pair, X^o and U^o, of optimal vectors satisfy the $2m + 2n$ feasibility constraints

$$AX \leqq B, \quad X \geqq 0 \quad \text{and} \quad UA \geqq C, \quad U \geqq 0.$$

There is a convenient way to pair off the constraints associated with the two dual problems. Let the i-th relations in the system $AX \leqq B$, $U \geqq 0$ be called a pair of <u>dual constraints</u>, as well as the j-th relations in the system $UA \geqq C$, $X \geqq 0$.

<u>Corollary 3.3.</u> If both the minimization and maximization problems have feasible vectors, then they have optimal vectors such that the dual constraints are complementary slack inequalities. That is, $m + n$ on the constraints are satisfied as equations and the remaining $m + n$ constraints are satisfied as strict inequalities.

Proof: Theorem 3.2 implies that both problems have optimal vectors. The last statement can be seen by substituting appropriately into the general dual system of Theorem 2.6. To relieve the problem of nomenclature, let the two systems of the Corollary be

(3.6) $\quad A'X \leqq B', \quad U' \geqq 0 \quad$ and $\quad X \geqq 0, \quad U'A' \geqq C'$.

Then in the general dual system of Theorem 2.6, since vectors U and Y are unrestricted, choose U, Y, A, and D such that $-U^TA = C'$ and $-DY = B'$. Let $V = (U')^T$ and $C = A'$. After making these substitutions, it is readily seen that the m + n pairs of inequalities referred to in Theorem 2.6 are those in (3.6). Hence, each of the m + n pairs of dual constraints contains exactly one inequality that is slack.

Corollary 3.3 may be stated more explicitly so that its implications will be readily available for use in later proofs. The restatement will require the following two corollaries.

<u>Corollary 3.4.</u> If both of the dual problems have feasible vectors, then they have optimal vectors X^o and U^o such that if X^o satisfies a row constraint as an equation, then U^o satisfies the dual constraint as a strict inequality; and if U^o satisfies a column constraint as an equation, then X^o satisfies the dual constraint as a strict inequality.

Proof: By Theorem 3.2, both problems have optimal vectors so that Lemma 3.2 implies that $t_o > 0$. Therefore, Lemma 3.1 implies that there exist optimal vectors X^o and U^o such that $(U^o)^T + B > AX^o$ and $U^oA + (X^o)^T > C$. The vectors thus exhibited have the required properties.

<u>Corollary 3.5.</u> If both of the dual problems have feasible vectors, then for any row index i, either $(AX^o)_i < b_i$ for some

optimal vector X^o and $u_i = 0$ for every optimal vector U, or $(AX)_i = b_i$ for every optimal vector X and $u_i^o > 0$ for some optimal vector U^o. The dual statement is also true.

Proof:. By Theorem 3.2, both of the dual problems have optimal vectors. Corollary 3.4 implies that if $(AX)_i = b_i$ for every optimal vector X, then $u_i^o > 0$ for some optimal vector U^o. However, if $(AX)_i \neq b_i$, then $(AX^o)_i < b_i$ for some optimal vector X^o. To complete the proof, it must be shown that $u_i = 0$ for every optimal vector U. Assume this not ture, i.e., $u_i > 0$ when $(AX^o)_i < b_i$ for some optimal vector X^o. This implies that $u_i(AX)_i < u_i b_i$ for some index i. Since X^o is feasible, $CX^o \leqq UAX^o$, or in particular, $(CX^o)_i \leqq u_i(AX^o)_i$. Hence, $(CX^o)_i < u_i b_i$. But, Corollary 3.1 implies that $(CX^o)_i = u_i b_i$ since X^o and U are optimal vectors. Thus, the contradiction implies that $u_i = 0$ for every optimal vector U.

Systems With Mixed Constraints

Linear programming problems involving mixed constraints (both equalities and inequalities) frequently arise in practice. A method due to A. W. Tucker (8), for reducing such a problem to canonical form is presented in this section.

Let N be the set of indices $(1,2,\ldots,n)$, and let M be the set of indices $(1,2,\ldots,m)$. Suppose that N_1 and N_2 are complementary subsets of N with n_1 and n_2 elements respectively, and that M_1 and M_2 are complementary subsets of M with m_1 and m_2 elements respectively. The dual linear programming problems with mixed constraints that will be investigated in this section may now be stated:

(4.1) Maximize $c_1x_1 + \cdots + c_nx_n$ subject to the constraints

$$a_{i1}x_1 + \cdots + a_{in}x_n \leqq b_i \quad \text{for each } i \text{ in } M_1,$$
$$a_{i1}x_1 + \cdots + a_{in}x_n = b_i \quad \text{for each } i \text{ in } M_2,$$
$$x_j \geqq 0 \quad \text{for each } j \text{ in } N_1,$$
$$x_j \text{ unrestricted for each } j \text{ in } N_2.$$

(4.2) Minimize $u_1b_1 + \cdots + u_nb_n$ subject to the constraints

$$u_1a_{1j} + \cdots + u_ma_{mj} \geqq c_j \quad \text{for each } j \text{ in } N_1,$$
$$u_1a_{1j} + \quad + u_ma_{mj} = c_j \quad \text{for each } j \text{ in } N_2,$$
$$u_i \geqq 0 \quad \text{for each } i \text{ in } M_1,$$
$$u_i \text{ unrestricted for each } i \text{ in } M_2 1$$

Note that the constraints which are equations correspond to variables which are unrestricted.

If the matrix A and the vectors B, C, X, and U are partitioned into clocks corresponding to the decomposition of M and N into $M_1 + M_2$ and $N_1 + N_2$ respectively, then, the problems (4.1) and (4.2) may be written in matrix form as follows:

Maximize $C_1X_1 + C_2X_2$ subject to the constraints

$A_{11}X_1 + A_{12}X_2 \leqq B_1$

$A_{21}X_1 + A_{22}X_2 = B_2$

$X_1 \geqq 0$, X_2 unrestricted.

Minimize $U_1B_1 + U_2B_2$ subject to the constraints

$U_1A_{11} + U_2A_{21} \geqq C_1$

$U_1A_{12} + U_2A_{22} = C_2$

$U_1 \geqq 0$, U_2 unrestricted.

The definition of feasible and optimal vectors, as previously given, apply as well to the more general problem in this section. That is, X or U is feasible if it satisfies the constraints and is optimal if it is feasible and achieves the desired maximum or minimum.

It should be emphasized that the problems (4.1) and (4.2) are essentially no more general than the problems (1.7) and (1.8), which are in canonical form. Indeed, (4.1) and (4.2) reduce to (1.7) and (1.8) respectively, if the sets M_2 and N_2 are vacuous. Furthermore, the method of replacing each equation of a system by two inequalities and each unrestricted variable of the system by the difference of two nonnegative variables results in problems in canonical form. However, the new maximization problem, for example, will have $m'=2m-\overline{M}_1$ inequalities and $n'=2n-\overline{N}_1$ nonnegative variables, where $\overline{M}_1$ denotes the number of elements in M_1 and $\overline{N}_1$ denotes the number of elements in N_1.

The method of the preceding paragraph can be implemented formally by first setting

$$(4.3)\qquad X_2'=\max(0,X_2) \qquad U_2'=\max(0,U_2)$$
$$X_2''=\max(0,-X_2) \qquad U_2''=\max(0,-U_2)$$

where $\max(Y^1, Y^2)$ is the vector whose components are $z_i=\max(y_i^1,y_i^2)$. From (4.3), it follows that $X_2=X_2'-X_2''$ and $U_2=U_2'-U_2''$. Secondly, each equation $E=0$ is replaced by the equivalent pair of inequalities, $E\geqq 0$ and $-E\geqq 0$. After performing these two steps, the problems (4.1) and (4.2) will be reduced to canonical form and may be stated as follows:

(4.4) Maximize $C_1X_1 + C_2X_2' + (-C_2)X_2''$ subject to the constraints

$$A_{11}X_1 + A_{12}X_2' + (-A_{12})X_2'' \leqq B_1$$
$$A_{21}X + A_{22}X_2' + (-A_{22})X_2'' \leqq B_2$$
$$(-A_{21})X_1 + (-A_{22})X_2' + A_{22}X_2'' \leqq (-B_2)$$
$$X_1\geqq 0,\quad X_2'\geqq 0,\quad X_2''\geqq 0.$$

(4.5) Minimize $U_1B_1 + U_2'B_2 + U_2''(-B_2)$ subject to the constraints

$$\begin{array}{llll} U_1A_{11} & + U_2'A_{21} & + U_2''(-A_{21}) & C_1 \\ U_1A_{12} & + U_2'A_{22} & + U_2''(-A_{22}) & C_2 \\ U_1(-A_{12}) & + U_2'(-A_{22}) & + U_2''A_{22} & (-C_2) \\ U_1 \quad 0, & U_2' \quad 0, & U_2'' \quad 0. & \end{array}$$

As previously noted, the dual problems (4.4) and (4.5) contain more constraints than the dual pair (4.1) and (4.2). However, the two problem pairs are equivalent in the sense that (U_1,U_2',U_2'') is feasible (optimal) if, and only if, (U_1,U_2) is feasible (optimal), and similarly for the dual problem. Therefore, the Duality and Existence Theorems of the last section are valid for the systems (4.1) and (4.2).

An algorithm for reducing a pair of linear programming problems with mixed constraints to canonical form will now be given. This method, mentioned in the opening paragraph of this section, does not have the disadvantage of the reduced canonical form containing more inequalities than the original problem. The method eliminates the equalities and unrestricted variables while the sum of the number of inequalities and the number of nonnegative variables remains constant under the reduction. The method utilizes the fact that any equation of the system containing a variable with non-zero coefficient can be used as a defining relation to eliminate that variable from the system.

Suppose, for example, that the m-th constraint is the equation $a_{m1}x_1 + \quad + a_{mn}x_n = b_m$, with $a_{mn} \quad 0$. That is m is in

M_2 and the variable U_m is unrestricted. Now, two situations can occur; either n is in N_1 or in N_2.

If n is in N_2, the n-th column constraint is an equation and the variable x_n is unrestricted. Then, the variables x_n and u_m can be eliminated from the systems by the relations

$$(4.6)\qquad x_n = \frac{1}{a_{mn}}(b_m - a_{m1}x_1 - \cdots - a_{m,n-1}x_{n-1}),$$
$$u_m = \frac{1}{a_{mn}}(c_j - u_1a_{1n} - \cdots - u_{m-1}a_{m-1,n}).$$

When these variables are eliminated, the elements a_{ij}, b_i, and c_j will be replaced by

$$(4.7)\qquad \bar{a}_{ij} = a_{ij} - \frac{a_{mj}a_{in}}{a_{mn}},$$
$$\bar{b}_i = b_i - \frac{b_m a_{in}}{a_{mn}},$$
$$\bar{c}_j = c_j - \frac{a_{mj}c_n}{a_{mn}}, \quad \text{for } i=1,\ldots,m-1;\ j=1,\ldots,n-1.$$

Note that the m-th row constraint and the n-th column constraint will also be eliminated.

The new dual problems, given by the matrics of lower degree which contain the elements (4.7), are equivalent to the old problems, (4.1) and (4.2), in the sense that $(x_1,\ldots,x_n)$ or $(u_1,\ldots,u_m)$ is feasible (optimal) for the old problem if, and only if, $(x_1,\ldots,x_{n-1})$ or $(u_1,\ldots,u_{m-1})$ is feasible (optimal) for the new problem.

If, however, n is in N_1, the n-th column constraint is an inequality and the variable x_n is nonnegative. Then, by making the same substitution for x_n as given by (4.6), replacing the m-th equation by

$$\frac{a_{m1}}{a_{mn}} x_1 + \cdots + \frac{a_{m,n-1}}{a_{mn}} x_{n-1} \leq \frac{b_m}{a_{mn}},$$

and the unrestricted variable u_m by the nonnegative variable

$$\bar{u}_m = u_1 a_{1n} + \cdots + u_m a_{mn} - c_n,$$

an equivalent problem is obtained. The maximization problem has one more inequality and one less nonnegative variable while in the minimization problem, an inequality has been replaced by demanding that another variable be nonnegative.

Again, the new dual problems obtained have feasible (optimal) vectors $(x_1, \ldots, x_{n-1})$ and $(u_1, \ldots, u_{m-1}, \bar{u}_m)$ if, and only if, $(x_1, \ldots, x_n)$ and $(u_1, \ldots, u_m)$ are feasible (optimal) for the old problems. In either case (n in N_1 or in N_2), the functions CX and UB are changed by the same constant $b_m c_n / a_{mn}$.

Evidently, a similar elimination reduces the number of equalities and unrestricted variables when applied to the column constraints that are equations. After a finite sequence of such operations, applied to both rows and column, one reaches a situation in which all constraint equations (if any exist) have only zero coefficients. If all of the constant terms (the b_i's and c_j's) in these equations are zero, then the associated rows and columns of zeros may be deleted. This yields a problem-pair in canonical form, i.e., $m_2 = n_2 = 0$.

If any of the constant terms are nonzero, then the corresponding member of the original problem-pair, is exhibited as unfeasible.

Systems of Equated Constraints

In the theory of linear programming, many significant results pertain to associated systems of the canonical form systems (1.7) and (1.8). The associated systems that will be discussed in this section contain only equations, namely, systems of equated constraints.

Let the systems of inequalities under consideration be the systems (1.7) and (1.8). Consider the various subsets of the $m + n$ inequalities in each system that are satisfied as equations. It will be convenient to describe these systems in the following manner. Let M_1 and M_2 be arbitrary subsets of the set M of indices $(1,\ldots,m)$, and let N_1 and N_2 be arbitrary subsets of the set N of indices $(1,\ldots,n)$. Then, each of the systems

$$(6.1)\qquad \begin{array}{ll} (AX)_i=b_i & \text{for all } i \text{ in } M_1, \\ x_j=0 & \text{for all } j \text{ in } N_2, \end{array}$$

and

$$(6.2)\qquad \begin{array}{ll} (UA)_j=c_j & \text{for all } j \text{ in } N_1, \\ u_i=0 & \text{for all } i \text{ in } M_2, \end{array}$$

is called a <u>system of equated constraints</u>.

With the maximization problem in mind, consider the set S of its feasible vectors X. The set S is defined by the finite system (1.7). It is the simple fact that an individual feasible vector X of S must satisfy an individual inequality $(AX)_i \leqq b_i$

of (1.7) either as an equation $(AX)_i=b_i$, or as a strict inequality $(AX)_i<b_i$, that gives structure to S. That is, each feasible vector in the set S may be classified according to which inequalities it satisfies strictly and which it satisfies as equations. This leads to a partitioning of the set S into 2^m disjoint subsets $S_{M'}$. The subsets $S_{M'}$ are indexed by means of the set M' of all indices i that specify the inequalities that are satisfied as equations. That is, the set $S_{M'}$ consists of all feasible vectors X that satisfy

$$(AX)_i=b_i \quad \text{for each } i \text{ in } M', \text{ and}$$
$$(AX)_i<b_i \quad \text{for each } i \text{ not in } M'.$$

It should be noted that some, or even all of the sets $S_{M'}$ may be vacuous. Of course, similar statements could be made about the dual problem.

The systems (6.1) and (6.2) are said to <u>dual</u> systems of equated constraints if M_1 and M_2 are complementary subsets of M and if N_1 and N_2 are complementary subsets of N. If the systems (6.1) and (6.2) are dual, the equations may be renumbered, if necessary, so that $M_1=(1,\ldots,p)$ and $N_1=(1,\ldots,q)$ and the systems written as

$$(6.3) \qquad \begin{array}{lcl} a_{11}x_1+\cdots+a_{1q}x_q=b_1 & & u_1a_{11}+\cdots+u_pa_{p1}=c_1 \\ \quad\ddots & \text{and} & \quad\ddots \\ a_{p1}x_1+\cdots+a_{pq}x_q=b_p & & u_1a_{1q}+\cdots+u_pa_{pq}=c_q \\ x_{q+1}=0,\ldots,x_n=0, & & u_{p+1}=0,\ldots,u_m=0. \end{array}$$

If a system of equated constraints, after being written in the form (6.3) is such that the coefficient matrix

$$A_1 = \begin{bmatrix} a_{11} & a_{12} \cdots a_{1q} \\ a_{21} & a_{22} \cdots a_{2q} \\ & \ddots \\ a_{p1} & a_{p2} \cdots a_{pq} \end{bmatrix}$$

is square and nonsingular, then the system is said to be a <u>nonsingular square system</u> of equated constraints.

<u>Theorem 5.1.</u> Feasible vectors are optimal if, and only if, they satisfy dual systems of equated constraints.

Proof: To prove the necessity of the condition, suppose that the vectors X and U are optimal. Define M_1 to be the subset of indices contained in M such that $(AX)_i=b_i$ and define M_2 to be the complementary subset of M, i.e., $M_2=M-M_1$. Then, by Corollary 3.5, the $u_i=0$ for all i in M_2. If N_1 and N_2 are defined analogously, then it follows from the same corollary that $x_j=0$ for all j in N_2. Therefore, the vectors X and U satisfy the dual systems of equated constraints indexed by M_1, M_2, N_1, and N_2.

To prove the sufficiency of the condition, suppose that feasible vectors X and U satisfy the dual systems (6.3). Then

$$UB=\sum^{p} u_i b_i = \sum^{p} u_i \Big(\sum^{q} a_{ij}x_j\Big) = \sum^{q}\Big(\sum^{p} u_i a_{ij}\Big)x_j = \sum^{q} c_j x_j = CX.$$

Thus, by Lemma 1.2, X and U are optimal vectors.

<u>Theorem 5.2.</u> There exists a unique pair of dual systems of equated constraints such that each is the maximal system of equated constraints satisfied by all optimal vectors.

Proof: Let M_1 denote the maximal set of row indices such that $(AX)_i = b_i$ for every optimal vector X. Similarly, let N_1 denote the maximal set of column indices such that $(UA)_j = c_j$ for all optimal vectors U. By Corollary 3.5, the row and column constraints indexed by elements of M_1 and N_1 respectively, determine dual systems of equated constraints which have the desired maximal property. It is also evident from Corollary 3.5 that the dual systems so determined are unique.

The concluding theorems deal with extreme solution vectors. An <u>extreme feasible</u> vector is a feasible vector which is not the mean, $\frac{1}{2}(X^1 + X^2)$ or $\frac{1}{2}(U^1 + U^2)$, of the two other feasible vectors. An <u>extreme optimal</u> vector is an optimal vector which is not the mean, $\frac{1}{2}(X^1 + X^2)$ or $\frac{1}{2}(U^1 + U^2)$, of the two other optimal vectors.

<u>Theorem 5.3.</u> A feasible vector X (or U) is an extreme feasible vector if, and only if, it satisfies a nonsingular square system of equated constraints.

Proof: To prove the sufficiency of the condition, suppose that a feasible vector X satisfies the nonsingular square system with the associated matrix A_1. Assume that X is not an extreme feasible vector, i.e., $X = \frac{1}{2}(X^1 + X^2)$, where X^1 and X^2 are feasible. Since, $x_j = 0$, $x_j^1 \geqq 0$, and $x_j^2 \geqq 0$ for

$j > p$, $x_j = \frac{1}{2}(x_j^1 + x_j^2) = 0$. Hence, $x_j^1 = x_j^2 = 0$ for $j > p$. Now, if $Y = (y_1, \ldots, y_p, \ldots, y_n)$ is any n-dimensional vector, define $\bar{Y} = (y_1, \ldots, y_p)$. Then, since X^1 and X^2 are feasible, $A_1\bar{X}^1 \leqq B_1$, $A_1\bar{X}^2 \leqq B_1$, and $\frac{1}{2}A(X^1 + X^2) = B_1$, where $B_1 = (b_1, \ldots, b_p)^T$. Therefore, $A_1\bar{X}^1 = A_1\bar{X}^2 = B_1$. Since A_1 is nonsingular, $\bar{X}^1 = \bar{X}^2$. Furthermore, $X^1 = X^2$, since the last n-p components of both X^1 and X^2 are zero. It then follows that $X = \frac{1}{2}(X^1 + X^2) = X^1 = X^2$ is not the mean of other feasible vectors. Hence, X is an extreme feasible vector of the given system.

To prove the necessity of the condition, a nonsingular square system of equated constraints satisfied by a given extreme feasible vector must be exhibited. Suppose that X is a nonzero extreme feasible vector of the system (1.7). Let M' denote the set of row indices for which $(AX)_i = b_i$. The set M' is not vacuous, for the assumption that M' is vacuous leads to a contradiction. The nonzero vector X has at least one component $x_j > 0$ and the vectors X^1 and X^2 obtained from X by replacing x_j with $x_j + \epsilon$ and $x_j - \epsilon$ respectively, have X as their mean. If M' were vacuous, i.e., if $(AX)_i < b_i$ for every i, then X^1 and X^2 would be feasible for sufficiently small $\epsilon > 0$. This contradicts the fact that X is an extreme feasible vector. Hence, M' is not vacuous.

Now, let N_2 be the set of column indices such that when j is in N_2, $x_j = 0$, and let N_1 be the complementary subset of N, the set of all column indices in the system (1.7). Note that N_1 is not vacuous because $X \neq 0$. Let $\bar{A}$ be the submatrix of the matrix A obtained by deleting from A the rows whose indices are not in M' and the columns whose indices are in N_2. Let $\overline{X}$ be the vector X with those zero components deleted whose indices are in N_2 and let $\overline{B}$ be the vector B with those components deleted whose indices are not in M'. Then clearly, $\overline{A}\overline{X}=\overline{B}$.

The linear independence of columns of $\overline{A}$ can now be easily established. If this were not true, there would exist a vector $\overline{X}^*\neq 0$ such that $\overline{A}\overline{X}^*=0$. Now, consider the vectors $\overline{X}_1=\overline{X}+\epsilon\overline{X}^*$ and $\overline{X}_2=\overline{X}-\epsilon\overline{X}^*$, which have dimension equal to n minus the number of indices in N_2. If zero components are adjoined to $\overline{X}_1$ and $\overline{X}_2$, to obtain n-dimensional vectors X_1 and X_2, the mean of X_1 and X_2 is X, while for sufficiently small $\epsilon>0$, X_1 and X_2 are feasible. This contradicts the fact that X is an extreme feasible vector. Hence, the columns of $\overline{A}$ are linearly independent.

The matrix $\overline{A}$ which has linearly independent columns is not necessarily square. However, the number of rows of $\overline{A}$ must be equal to or greater than the number of columns of $\overline{A}$. This can be seen by noting that a set of $\overline{N}_1$-dimensional vectors is linearly dependent if the set contains more than $\overline{N}_1$ vectors. Thus, by deleting a suitable number of rows of the matrix $\overline{A}$, a square nonsingular submatrix A_1 is obtained. Since, $\overline{A}\overline{X}=\overline{B}$, clearly, the extreme feasible vector X satisfies the system of equated constraints whose associated matrix is A_1.

In the above proof of the necessity of the condition, the vector was assumed to be nonzero. If, however, $X=0$, it is clear that X is an extreme feasible vector, if it is feasible. The parenthetical assertion of the Theorem can be proved in a similar manner.

Theorem 5.4. Feasible nonzero vectors X and U are extreme optimal vectors if, and only if, they satisfy dual nonsingular square systems of equated constraints.

Proof: The sufficiency of the condition follows immediately from Theorems 5.1 and 5.3. To prove the necessity of the condition, suppose that X and U are extreme optimal vectors. Let the sets of indices M_1, M_2, N_1, and N_2 be renumbered, if necessary, so that $(AX)_i=b_i$ for i in $M_1=(1,\ldots,p)$, $(AX)_i< b_i$ for i in $M_2=(p+1,\ldots,m)$, $(UA)_j=c_j$ for j in $N_1=(1,\ldots,q)$, and $(UA)_j> c_j$ for j in $N_2=(q+1,\ldots,n)$. That the sets M_1 and N_1 are not vacuous follows from the same argument given in the proof of Theorem 5.3. By Corollary 3.5, $x_j=0$ for $j>q$ and $u_i=0$ for $i>p$. However, some of the x_j with $j\leqq q$ or some of the u_i with $i\leqq p$ may also equal zero; in this case, let the sets of indices be renumbered again so that $x_j>0$ for $1\leqq j\leqq \bar{q}\leqq q$, $x_j=0$ for $j>\bar{q}$, $u_i>0$ for $1\leqq i\leqq \bar{p}\leqq p$, and $u_i=0$ for $i>\bar{p}$. Since $X\neq 0$ and $U\neq 0$, $\bar{p}>0$ and $\bar{q}>0$, and hence, $p>0$ and $q>0$.

Let $\bar{A}$ denote the submitrix of the matrix A in (1.7) which contains as elements the intersection of the first p rows and the first q columns of A.

The linear independence of the first $\bar{q}$ columns of $\bar{A}$

can be established using an argument similar to the one used in the proof of Theorem 5.3. If the first $\bar{q}$ columns of $\bar{A}$ are linearly dependent, there must exist a $\bar{q}$-dimensional vector $\bar{X}^{*} \neq 0$ such that $\bar{A}\bar{X}^{*}=0$. Consider the vectors $\bar{X}_1=\bar{X}+\epsilon\bar{X}^{*}$ and $\bar{X}_2=\bar{X}-\epsilon\bar{X}^{*}$ where $\bar{X}$ is the vector with the last $n - q$ zero components deleted. If zero components are adjoined to $\bar{X}_1$ and $\bar{X}_2$ to obtain n-dimensional vectors X_1 and X_2, the mean of X_1 and X_2 is X while for sufficiently small $\epsilon > 0$, X_1 and X_2 are feasible. The fact that the vectors X_1 and X_2 are also optimal follows readily from the relations $CX_1 \leqq UB$, $CX_2 \leqq UB$, and $CX=\frac{1}{2}C(X_1 + X_2)=UB$, which are valid because X_1 and X_2 are feasible and X is optimal. Indeed, $CX_1 + UB \geqq CX_1 + CX_2 = 2UB$ or $CX_1 \geqq UB$. Hence, $CX_1=UB$, and similarly, $CX_2=UB$. By Lemma 1.2, the vectors X_1 and X_2 are optimal. Thus, the fact that the vectors X_1 and X_2 are optimal with mean X contradicts the fact that X is an extreme optimal vector and leads one to the conclusion that the first $\bar{q}$ columns of $\bar{A}$ are linearly independent. Similarly, the first $\bar{p}$ rows of $\bar{A}$ are linearly independent.

Now, consider the maximal set of rows and the maximal set of columns of $\bar{A}$. Let the sets of indices indicating these rows and columns be renumbered so that the first p' rows of $\bar{A}$ $(p' \geqq p)$ are linearly independent and the first q' columns of $\bar{A}$ $(q' \;\; q)$ are linearly independent. Denote the submatrix which contains as elements, the intersection of the first p' rows and the first q' columns of $\bar{A}$ by A_1. The last $q - q'$

columns of $\bar{A}$ are linearly dependent on the first q' columns and, therefore, any linear combination of the p' rows of A_1 can be extended to a linear combination of the p' rows of $\bar{A}$ by virtue of this linear dependence. Now, the p' rows of $\bar{A}$ are linearly independent. Therefore, the rows of A_1, which contain only parts of the components of linearly independent vectors, must be linearly independent. Similarly, the columns of A_1 are linearly independent. Hence, A_1 is square and nonsingular. Clearly, the vectors X and U satisfy the dual nonsingular square system of equated constraints with associated matrix A_1.

Corollary 5.1. The sets of feasible and optimal vectors of dual linear programs have a finite number of extreme vectors.

Proof: This statement follows immediately from Theorem 5.3 and 5.4.

It is easily seen that Theorem 5.4 provides a method for finding extreme optimal vectors. However, to use this method, one must determine all square systems of equated constraints with nonsingular associated matrices and find the feasible solutions of these systems. This procedure would, in general, require a tremendous amount of work.

Optimal Rays

An optimal ray $(X^0;X)$ for the maximization problem (1.7) is a set of optimal vectors of the form $X^0 + \lambda X$, where λ takes on all nonnegative values, X^0 is a fixed optimal vector, and X is any fixed vector which has been normalized so that the sum of its components is one. Optimal rays for the minimization problem are defined analogously. The vectors, X, are called directions of the optimal rays and the following Lemma characterizes them.

Lemma 6.1. Suppose the sum of the components of the vector X is one, and that X^0 is an optimal vector for the maximization problem (1.7). Then, (X^0,X) is an optimal ray if, and only if,

(1) $X \geqq 0$

(2) $AX \leqq 0$

(3) $CX=0$.

The dual statement holds as well; (2) becomes $UA \geqq 0$.

Proof: The set of optimal vectors $(X^0;X)$ is an optimal ray if, and only if,

(i) $AX^0 + \lambda AX \leqq B$

(ii) $CX^0 + \lambda CX = CX^0$

(iii) $X^0 + \lambda X \geqq 0$ for all $\lambda \geqq 0$.

Clearly, these conditions are satisfied if, and only if, (1), (2), and (3) hold.

Lemma 6.2. The set of directions of optimal rays of a linear-programming problem has a finite number of extreme vectors.

Proof: According to Lemma 6.1, $\bar{X}$ 0 is the direction of an optimal ray for the maximization problem if, and only if,

(i) $\sum x_j \leqq 1$

(ii) $\sum (-x_j) \leqq 1$

(iii) $AX \leqq 0$

(iv) $CX \leqq 0$

(v) $(-CX) \leqq 0$.

Relations (i) through (v) can be considered the row constraints of a suitable new maximization problem. The desired conclusion follows by applying Corollary 5.1 to this problem. Of course, an analogous argument holds for the minimization problem.

The discussion of the theory necessary to prove the following theorems is quite lengthy and involves concepts which have not been considered in this report. For this reason, these theorems will be stated without proof. Theorem 6.1 is needed to prove Theorem 6.2, a statement which characterizes the set of all optimal vectors of a linear-programming problem. A proof of Theorem 6.1 can be found in (7) and a proof of Theorem 6.2 can be found in (8).

Theorem 6.1. Let $S = \{X \mid AX \leqq B\}$ be a nonvacuous set with the matrix A having linearly independent columns. Then S has a minimal basis $(P_1, \ldots, P_p; Q_1, \ldots, Q_q)$ which is unique (up to positive multiples of the Q_j's). Here, $(P_1,\ldots,P_p)$ is the set of extreme vectors of S, and $(Q_1,\ldots,Q_q)$ is the set of directions of the extreme vectors of S.

<u>Theorem 6.2.</u> Let $\{X^r\}$ be the finite set of extreme optimal vectors for the maximization problem and $\{X^s\}$ be the finite set of extreme directions of optimal rays for the maximization problem. Then, the set of all optimal vectors X is the set of all vectors of the form $X=\sum \lambda_r X^r + \sum \mu_s X^s$ with all $\lambda_r \geqq 0$, all $\mu_s \geqq 0$, and $\sum \lambda_r = 1$. Of course, the dual statement holds for the minimization problem.

Lagrange Multipliers

The traditional calculus procedure for finding a constrained maximum or minimum employs Lagrange Multipliers. This procedure may be adapted to linear programming problems to derive a necessary and sufficient condition for the vectors X and U to be optimal vectors for the dual programs (1.7) and (1.8).

The general procedure of Lagrange's method will now be summarized. If a given function $G(x_1,\ldots,x_n)=G(X)$ is to be maximized or minimized subject to the constraints, $F_1(X)=0,\ldots,F_m(X)=0$, then the first step is to form the Lagrangian function $H(X,u_1,\ldots,u_m)=H(X,U)=G(X) + \sum_{i=1}^{m} u_i F_i(X)$. Then, the necessary conditions that $H(X,U)$ have an unconstrained extremum are that the n first partial derivations of $H(X,U)$ vanish. These conditions are also necessary for $G(X,U)$ to have a constrained extremum, so that the free extreme values of $H(X,U)$ are sought, among which will be the extreme values of $G(X,U)$.

A similar situation arises in linear programming, where the object is to extremize a linear function subject to constraints which are linear inequalities, rather than equations. To develop a technique analogous to Lagrange's method, the Lagrangian function

$$\begin{aligned} (7.1) \qquad L(X,U) &= CX + UB - UAX \\ &= CX + \sum_i u_i(b_i - \sum_j a_{ij}x_j) \\ &= UB + \sum_j x_j(c_j - \sum_i u_i a_{ij}), \end{aligned}$$

is formed. The second form of $L(X,U)$ in (7.1) exhibits a situation analogous to that described above. The problem described here is a maximization problem with $G(X) = CX$, $F_i(X)=(b_i - \sum_j a_{ij}x_j)$, and the u_i regarded as multipliers. Similarly, the third form of $L(X,U)$ in (7.1) exhibits the analogy for the minimization problem. The following theorem justifies this analogy.

Theorem 7.1. A necessary and sufficient condition that $X^o \geqq 0$ and $U^o \geqq 0$ be optimal vectors for the dual programs (1.7) and (1.8) is that (X^o, U^o) be a "saddle point" for $L(X,U)$ in the sense that

$$L(X,U^o) \leqq L(X^o,U^o) \leqq L(X^o,U)$$

for all $X \geqq 0$ and $U \geqq 0$. If X^o and U^o are optimal vectors, then $L(X^o,U^o)$ is the optimal value of the dual programs.

Proof: To prove the necessity of the condition, suppose that X^o and U^o are optimal vectors. By Corollary 3.1, $CX^o=U^oB$. Since $CX^o \leqq U^oAX^o \leqq U^oB$, it follows from (7.1) that $L(X^o,U^o) = CX^o = U^oB$. Thus, for any $X \geqq 0$, $U \geqq 0$,

$$L(X,U^o) = U^oB + (C - U^oA)X \leqq U^oB = L(X^o,U^o) \quad \text{and}$$
$$L(X^o,U) = CX^o + U(B - AX^o) \geqq CX^o = L(X^o,U^o).$$

Therefore, the point (X^o,U^o) is a saddle point.

To prove the sufficiency of the condition, suppose that (X^o,U^o) is a saddle point with $X^o \geqq 0$ and $U^o \geqq 0$. Then, for any $X \geqq 0$,

$$U^oB - (U^oA - C)X = L(X,U^o) \leqq L(X^o,U^o) = U^oB - (U^oA - C)X^o$$

or

$$(7.2) \qquad (U^oA - C)(X - X^o) \geqq 0.$$

Now, since X is arbitrary, choose $X = X^o + E_j$, where E_j is the j-th unit vector. Then, (7.2) yields $(U^oA - C)_j \geqq 0$. Let $j = 1,\ldots,n$, and it follows that $U^oA \geqq C$. Hence, U^o is feasible. Similarly, for any vector $U \geqq 0$,

$$CX^o + U^o(B - AX^o) = L(X^o,U^o) \leqq L(X^o,U) = CX^o + U(B - AX^o)$$

or

$$0 \leqq (U - U^o)(B - AX^o).$$

Therefore, choose $U = U^o + E_i$; let $i=1,\ldots,m$, and it follows, as above, that $B \geqq AX^o$. Hence, X^o is feasible.

Since $L(0,U^o) \leqq L(X^o,U^o) \leqq L(X^o,0)$, $U^oB \leqq CX^o$. By Lemma 1.1, $CX^o \leqq U^oB$. Hence, $U^oB=CX^o$ and by Lemma 1.2, vectors X^o and U^o are optimal. Therefore, $U^oB=L(X^o,U^o)=CX^o$.

It is to be noted that the situation involving dual linear programs to which the method of Lagrange has been applied here differs from the ordinary situation in the calculus. First, the fact that a maximization problem and a minimization problem are being dealt with simultaneously leads to the consideration of a saddle point of $L(X,U)$. Second, the conditions involved are both necessary and sufficient. Finally, the Lagrangian function $L(X,U)$ is subject to the constraints $X \geqq 0$ and $U \geqq 0$.

Appendix

One of the best known and the most widely used computational methods of solving a linear programming problem is the Simplex Method developed by G. B. Dantzig (3). The purpose of this section is to discuss this method.

The linear-programming problem to be discussed in this section will considered in the following standard form:

Find the values $\lambda_1, \ldots, \lambda_n$ which maximize the linear function

(1) $$z_o = \lambda_1 c_1 + \cdots + \lambda_n c_n .$$

subject to the conditions that

(2) $$\lambda_j \geqq 0 \qquad (j=1,\ldots,n)$$

and

(3) $$\begin{aligned} \lambda_1 a_{11} + \cdots + \lambda_n a_{1n} &= b_1, \\ &\ddots \\ \lambda_1 a_{m1} + \cdots + \lambda_n a_{mn} &= b_m, \end{aligned}$$

where a_{ij}, b_i, c_j are constants $(i=1,\ldots,m;\ j=1,\ldots,n)$.

The general linear-programming problem may involve constraints which are inequalities or a mixture of equations and inequalities, and variables which can have negative values, and a linear function which is to be minimized rather than maximized. Such problems are easily transformed to the standard form stated above.

If a constraint is a less-than condition, so that $a_{i1}x_1 + \cdots + a_{in}x_n \leqq b_i$, then the constraint may be transformed into an equation by adding a nonnegative <u>slack</u> variable s_i to the left-hand side and writing $a_{i1}x_1 + \cdots + a_{1n}x_n + s_i = b_i$. The slack variable s_i is an additional unknown that has to be determined. Similarly, if a constraint is a greater-than condition, so that $a_{i1} + \cdots + a_{in}x_n \geqq b_i$, then one can write $a_{i1}x_1 + \cdots + a_{in}x_n - s_j = b_i$, where s_i is again a nonnegative slack variable. By introducing a nonnegative slack variable for each constraint that is an inequality, one can always express the constraints in the standard form of (3).

If some variables are not constrained to be nonnegative, then those variables can always be expressed as the difference of two nonnegative variables. Substituting the difference of two nonnegative variables into (3) for those variables which are not restricted in sign will result in a problem in standard form.

Finally, if one wishes to minimize the function (1), rather than maximize it, one may reverse the signs of all the c_j in (1) and then proceed to maximize the resulting function.

Each column of coefficients in (3) may be viewed as representing the components of a vector in the vector space V_m. Let P_j denote the j-th column of coefficients and P_o the column of constants in (3). Then, by definition,

(4)

$$(P_1,\ldots,P_n;\ P_o)=\begin{bmatrix} a_{11} a_{12} \cdots a_{1n} b_1 \\ \cdot \\ \cdot \\ \cdot \\ a_{m1} a_{m2} \cdots a_{mn} b_m \end{bmatrix}.$$

The basic problem then is to determine nonegative $\lambda_j \geqq 0$ such that

(5) $$\lambda_1 P_1 + \cdots + \lambda_n P_n = P_o,$$

(6) $$\lambda_1 c_1 + \cdots + \lambda_n c_n = z = \max.$$

A set of λ_i which satisfy (5) without necessarily yielding the maximum in (6) will be termed a feasible solution; one which maximizes (6) will be called a maximum feasible solution. Having defined the problem, the simplex method for finding its solutions may now be discussed. The simplex technique consists of constructing first a feasible, and then a maximum feasible, solution. Since feasible solutions are frequently obtained by inspection and because an arbitrary feasible solution can be obtained in a manner analogous to the construction of a maximum feasible solution, the construction of a maximum feasible solution from a given feasible solution will be considered first.

To simplify the development that follows, the following nondegeneracy assumption will be made: Every subset of m vectors from the set $(P_o;P_1,\ldots,P_n)$ is linearly independent.

However, modifications to the method have been developed that will solve any problem they may arise which is degenerate.

Construction of a Maximum Feasible Solution

Assume as given a feasible solution consisting of exactly m of the λ_i nonzero, that is,

(7) $$\lambda_1 P_1 + \cdots + \lambda_m P_m = P_o, \qquad \lambda_i > 0.$$

(8) $$\lambda_1 c_1 + \cdots + \lambda_m c_m = z_o.$$

The set of m vectors, P_j, in (7) will be referred to as a basis for $P_1, \ldots, P_n$. In establishing the conditions for the construction of a maximum feasible solution, it will be necessary to express all n vectors, P_j, in terms of a basis, that is,

(9) $$x_{1j} P_1 + \cdots + x_{mj} P_m = P_j \qquad (j=1,\ldots,n).$$

Now, define z_j by

(10) $$x_{1j} c_1 + \cdots + x_{mj} c_m = z_j \qquad (j=1,\ldots,n).$$

Theorem 1. If, for any fixed j, the condition

(11) $$c_j > z_j$$

holds, then a set of feasible solutions can be constructed such that

(12) $$z > z_o$$

for any member of the set, where the upper bound of z is either finite or infinite.

Case I: If finite, a feasible solution consisting of exactly m $\lambda_i > 0$ can be constructed.

Case II: If infinite, a class of feasible solutions consisting of a set of exactly $(m + 1)$ positive λ_i can be constructed such that the upper bound of $z=+\infty$.

Proof: Multiplying equation (9) by θ and substracting from equation (7), and multiplying (10) by θ and substracting from (8), the equations

(13) $$(\lambda_1 - \theta x_{1j})P_1 + \cdots + (\lambda_m - \theta x_{mj})P_m + \theta P_j = P_o,$$

(14) $$(\lambda_1 - \theta x_{1j})c_1 + \cdots + (\lambda_m - \theta x_{mj})c_m + \theta c_j = z_o + \theta(c_j - z_j),$$

are obtained where the term θc_j has been added to both sides of (14).

Since $\lambda_i \geqq 0$ for every i in (13), it is clear that there is, for $\theta \geqq 0$, either a finite range of values $\theta_o > 0 \geqq 0$ or an infinite range of values such that the coefficients of the P_i remain positive. It follows from (14) that the function z of this set of feasible solutions is a strictly monotonically increasing function of θ,

(15) $$z = z_o + \theta(c_j - z_j) > z_o, \qquad \theta > 0,$$

since $c_j > z_j$ by hypothesis (11). This establishes (12).

Case I: If $x_{ij} > 0$ for at least one $i=1,\ldots,m$ in (9) or (13), the largest value of θ for which all coefficients in (13) remain nonnegative is given by

(16) $$\theta_o = \min_i(\lambda_i / x_{ij}), \qquad x_{ij} > 0.$$

If for some value i_o of i, θ in (16) exists, it is clear that the coefficient corresponding to i_o in (13) and (14) will vanish. Hence, a feasible solution, given by $\theta=\theta_o$, has been constructed with exactly m positive weights λ_i. Furthermore, $z > z_o$. It should be noted that this new set of m vectors consists of the new vector, P_j, and $(m - 1)$ of the original m vectors. This is a desired solution for Case I.

The new set of m vectors may now be used as a new basis, and again, as in (9) and (10), all vectors may be expressed in terms of the new basis and the values of the c_j compared with the values of the z_j just computed. If any $c_j > z_j$, the value of z can be increased. If at least one $x_{ij} > 0$, another new basis can be formed. Now, assume that this process is continued until it is not possible to form a new basis. This must occur in a finite number of steps because there are at most ${}_nC_m$ bases and none of these can recur, for in that case the value of z would also recur. This cannot happen because the function z is strictly increasing. Thus, it is evident that the process must terminate, either because at some stage

(17) $$x_{ij} \leqq 0 \quad \text{for all} \quad i=1,\ldots,m \quad \text{and some fixed} \quad j,$$

or because

(18) $$c_j \leqq z_j \quad \text{for all} \quad j=1,\ldots,n.$$

Case II: If inequality (17) occurs, then it is clear that θ has no finite upper bound and that a class of feasible solutions has been constructed consisting of a set of $(m + 1)$ positive λ_i such that the upper bound of $z = +\infty$.

In all problems in which there is a finite upper bound to z, the iterative process must necessarily lead to the condition (18). However, it will be proved in the next theorem that the feasible solution associated with the final basis, which has the property that $c_j \leqq z_j$ for all $j=1,\ldots,m$ is also a maximum feasible solution. Hence, in all problems in which there is no finite upper bound to z, the iterative process must necessarily lead to the condition (17). Furthermore, by rewriting equation (9) as

$$P_j + (-x_{1j})P_1 + \cdots + (-x_{mj})P_m = 0, \qquad x_{ij} \leqq 0,$$

for the fixed j of (17), a nonnegative linear combination of $(m + 1)$ vectors has been shown to vanish if the upper bound of z is $+\infty$.

As a practical computing matter, the procedure of progressing from one basis to the next does not involve as much computation as would first appear, because the basis, except for the deletion and insertion of one vector, is the same as before. In fact, one iteration involves less than mn multiplications and an equal number of additions. It has been observed from experience that the number of iterations can be greatly reduced not by arbitrarily selecting any vector, P_j, satisfying $c_j > z_j$, but by selecting the one which gives the greatest immediate increase to the function z. From (15), the criterion for the choice of

j is such that $\theta_o(c_j - z_j)$ is a maximum, where θ_o is given by (16). A criterion that involves considerably less computation and seems to yield just as satisfactory results is to choose j such that $(c_j - z_j)$ is a maximum. In either case, approximately m changes in basis are encountered in practice, so that about m^2n multiplications are involved in obtaining a maximum feasible solution from a feasible solution.

<u>Theorem 2.</u> If, for all $j=1,\ldots,n$, the condition $c_j \leqq z_j$ holds, then equations (7) and (8) constitute a maximum feasible solution.

Proof: Let

(19) $$\mu_1 P_1 + \cdots + \mu_n P_n = P_o, \quad \mu_j \geqq 0,$$

(20) $$\mu_1 c_1 + \cdots + \mu_n c_n = z^*,$$

constitute any other feasible solution. It will be shown that $z_o \geqq z^*$.

By hypothesis, $c_j \leqq z_j$, so that replacing c_j by z_j in (20) gives

(21) $$\mu_1 z_1 + \quad + \mu_n z_n \geqq z^*.$$

On substituting the value of P_j given by equation (9) into equation (19) and the value of z_j given by (10) into (21), it follows that

(22) $$\left(\sum_{j=1}^{n} \mu_j x_{1j}\right) P_1 + \cdots + \left(\sum_{j=1}^{n} \mu_j x_{mj}\right) P_m = P_o,$$

(23) $$\left(\sum_{j=1}^{n} \mu_j x_{1j}\right)c_1 + \cdots + \left(\sum_{j=1}^{n} \mu_j x_{mj}\right)c_m \geqq z^*.$$

According to the assumption of nondegeneracy, the corresponding coefficients of P_i in equations (7) and (22) must be equal. Hence, the inequality (23) becomes $\lambda_1 c_1 + \cdots + \lambda_m c_m \geqq z^*$, or, by (8), $z_0 \geqq z^*$.

In order that another maximum feasible solution exist, it is necessary that $c_j = z_j$ for some P_j not in the final basis. However, in this case,

(24) $$\begin{bmatrix} P_1 & P_2 & P_n \\ c_1 & c_2 & c_n \end{bmatrix}.$$

has at least one set of $(m + 1)$ columns which are linearly dependent. Thus, a sufficient condition that the maximum feasible solution constructed from the given feasible solution be unique is that ever set of $(m + 1)$ vectors defined by columns in (24), be linearly independent.

Construction of a Feasible Solution

Consider the $(n + 1)$ column vectors in (4). Select an arbitrary basis of $(m - 1)$ vectors, P_j, and the vector P_o from this set. Denote the basis by $(P_o;P_1,\ldots,P_{m1})$. Now, any P_j can be expressed in terms of this basis by

(24) $$y_{oj}P_o + y_{1j}P_1 + \cdots + y_{(m-1)j}P_{m-1} = P_j \quad (j=1,\ldots,n).$$

<u>Theorem 3.</u> A sufficient condition that there exist no feasible solution is that $y_{oj} \leqq 0$ for all j.

Proof: Assume that there does exist a feasible solution,

(25) $\lambda_1 P_1 + \cdots + \lambda_n P_n = P_o, \quad \lambda_j \geqq 0.$

On substituting the expressions for P_j given by (24) into equation (25), it follows that

(26) $$P_o\Big(\sum_{j=1}^{n} \lambda_j y_{oj} - 1\Big) + P_1\Big(\sum_{j=1}^{n} \lambda_j y_{ij}\Big) + \cdots + P_{m-1}\Big(\sum_{j=1}^{n} \lambda_j y_{(m-1)j}\Big) = 0.$$

By virtue of the assumed independence of $(P_o; P_1, \ldots, P_{m-1})$, it is evident that each coefficient in (26) must vanish. In particular, $\sum_{j=1}^{n} \lambda_j y_{oj} - 1 = 0$. This is impossible if both $\lambda_j \geqq 0$ and $y_{oj} \leqq 0$ for all j.

To construct a feasible solution, a fixed reference vector R is first defined. R is given by $R = w_1 P_1 + \cdots + w_{m-1} P_{m-1} - k_o P_o$, where $w_i > 0$ $(i=1,\ldots,m-1)$ and $k_o > 0$ are arbitrarily chosen. The above equation may be rewritten in the form

(27) $R + k_o P_o = w_1 P_1 + \cdots + w_{m-1} P_{m-1}.$

In the development that follows, k_o will play a role analogous to z_o.

By Theorem 3, if there exists a feasible solution, there exists at least one j (which will be considered to fixed) such that $y_{oj} > \theta$. Multiplying (24) by θ and substracting from (27), it follows that

$$R + (k_o + \theta y_{oj}) P_o = \theta P_j + (w_1 - \theta y_{1j}) P_1 + \cdots + (w_{m-1} - \theta y_{(m-1)} P_{m-1}.$$

For a range of $\theta_o > \theta > 0$, one can construct, in a manner similar to (13) and (14), a set of vectors of the form $R + kP_o$, where each is a positive linear combination of the vectors P_j. Since k will play a role analogous to z, the problem is to find the largest value k for which this construction is possible. It should be noted that $k=k_o + \theta y_{oj} > k_o$ since $y_{oj} > 0$ has been assumed.

If, in the equation (24), all $y_{ij} \leqq 0$ $(i=1,\ldots,m-1)$, the coefficients of P_j will be positive and $k \rightarrow +\infty$ as $\theta \rightarrow +\infty$. Also, it will be seen, by solving (24) for P_o, that

$$P_o=(1/y_{oj})P_j + (-y_{1j})P_1 + \cdots + (-y_{(m-1)j}/y_{oj})P_{m-1},$$

and that a feasible solution has been obtained. That is, P_o has been expressed as a positive linear combination of $P_1,\ldots,P_{m-1}$ and P_j. If at least one $y_{ij} > 0$ $(i=1,\ldots,m-1)$, the largest value of θ is then

$$\theta_o=\min(w_i/y_{ij}), \qquad y_{ij} > 0. \tag{28}$$

When θ is set equal to θ_o, the coefficient of at least one vector, P_i, will vanish and a new vector $R + k_1P_o$, will be formed from (27) which is expressed as a positive linear combination of just $(m - 1)$ vectors, P_i, where $k_1=k_o + \theta_o y_{oj} > k_o$.

By expressing all vectors P_j in terms of the new basis, the process may be repeated, each time obtaining a larger value of k, i.e., a feasible solution. The process must terminate in a finite number of steps. Otherwise, since there is only a finite number of bases, the same combination of $(m - 1)$ points

P_i would appear twice. That is,

$$(29) \qquad R + k'P_o = w_1'P_1 + \cdots + w_{m-1}'P_{m-1},$$

$$(30) \qquad R + k''P_o = w_1''P_1 + \cdots + w_{m-1}''P_{m-1},$$

where $k'' > k'$. Subtracting equation (29) from equation (30), a nonvanishing expression giving P_o in terms of $(m - 1)$ vectors P_i, is obtained. This contradicts the nondegeneracy assumption.

There are only two conditions which will terminate the process. That is, after a finite number of iterations, either $y_{oj} \leqq 0$ for all $j=1,\ldots,n$, or, for some fixed j, $y_{ij} \leqq 0$ for all $i=1,\ldots,m$. In the first case, Theorem 3 implies that no feasible solution exists. In the latter case, equation (24) may be solved for P_o, as was done in (28), to obtain the desired feasible solution.

ACKNOWLEDGMENT

The writer wishes to express his sincere appreciation to Dr. S. Thomas Parker for his patient guidance and supervision given during the preparation of this report.

REFERENCES

1. Cooper, W. W., and A. Charnes. An Introduction to Linear Programming. New York: John Wiley & Sons, 1953. 74p.

2. Churchman, C. W., and R. L. Ackoff, and E. L. Arnoff. Introduction to Operations Research, Chapter 11. New York: John Wiley & Sons, 1961. pp. 279-341.

3. Dantzig, George B. "Maximization of a Linear Function of Variables Subject to Linear Inequalities," Activity Analysis of Production, Cowles Commission Monograph No. 13. New York: John Wiley & Sons, 1951. pp 339-347.

4. Ferguson, Robert O., and Lauren F. Sargent. Linear Programming: Fundamentals and Applications. New York: McGraw-Hill, 1958. 342p.

5. Ficken, Frederick A. The Simplex Method of Linear Programming. New York: Holt, Rinehart and Winston, 1961. 58p.

6. Garvin, Walter W. Introduction to Linear Programming. New York: McGraw-Hill, 1960. 281p.

7. Goldman, A. J. "Resolution and Separation Theorems for Polyhedral Sets," Annals of Mathematics Studies, No. 38. Princeton: Princeton University Press, 1956. pp. 19-40.

8. Goldman, A. J., and A. W. Tucker. "Theory of Linear Programming," Annals of Mathematics Studies, No. 38. Princeton: Princeton University Press, 1956. pp. 53-80.

9. Good, R. A. "Systems of Linear Relations," SIAM Review, Vol. 1, January, 1959. pp. 1-30.

10. Heady, Earl O., and Wilfred Candler. Linear Programming Methods. Ames Iowa State College Press, 1958. 597p.

11. Tucker, A. W.
"Dual Systems of Homogeneous Linear Relations," <u>Annals of Mathematics Studies</u>, No. 38. Princeton: Princeton University Press, 1956. pp. 3-19.

12. Tucker, A. W.
"Linear Inequalities and convex Polyhedral Sets," <u>Proceedings of the Second Symposium in Linear Programming</u>. Washington, D. C.: National Bureau of Standards, January, 1955. pp. 569-602.

THE THEORY OF LINEAR PROGRAMMING

by

ELMER EARL BRANSTETTER

B. S., Central Missouri State College, 1961

AN ABSTRACT OF A REPORT

submitted in partial fulfillment of the

requirements for the degree

MASTER OF SCIENCE

Department of Mathematics

KANSAS STATE UNIVERSITY
Manhattan, Kansas

1963

Linear programming is usually thought of as referring to techniques for solving a general class of optimization problems dealing with the interaction of many variables subject to certain linear restraining conditions. In solving these problems, certain objectives are to be obtained in the best possible, or optimal, fashion subject to certain restraining conditions which may arise from a variety of sources. However, the phases of linear programming discussed in this report are primarily existence discussions. The linear-programming problem is first defined in terms of a system of linear inequalities, and then, certain conditions necessary for the existence of a solution are derived.

The concept of duality is employed whenever possible. If the linear-programming problem is stated in terms of seeking the maximum value of a linear function subject to a system of linear constraints, then the dual to this problem is the search for a minimum value for an associated linear function subject to an associated system of linear constraints.

As background necessary to prove the fundamental Duality and Existence Theorems, dual systems of homogeneous linear relations are first discussed. These theorems give necessary and sufficient conditions for the existence of feasible and optimal solutions for the dual linear-programming problems.

Dual linear programs involving systems of mixed constraints are shown to be equivalent to the dual linear programs as originally defined.

From a study of linear programs involving systems of equated constraints, the set of all optimal solutions is characterized.

A modified version of the method of Lagrange Multipliers is used to derive a necessary and sufficient condition for the existence of optimal solutions for the dual linear programs.

In the appendix, the so-called simplex method, developed by George B. Dantzig, is discussed. It was the inability to find an analytic solution to the linear-programming problem that led to the development of the simplex method and various other iterative methods. The simplex method is the most widely used method, since experience has shown it to be a rather efficient method from a computational point of view.

While it is true that the theory of convex polyhedral cones is but a geometric interpretation of the theory of linear inequalities, this theory is quite extensive, and as a result is not emphasized in this report. Therefore, no attempt is made at geometrical interpretation of any concept. In fact, all results are stated and proved algebraically, as properties of linear inequalities.