1976

Computational aspects of linear multiple objective optimization

John Mark Trzeciak
Iowa State University

77-15

TRZECIAK, John Mark, 1949-
COMPUTATIONAL ASPECTS OF LINEAR MULTIPLE OBJECTIVE OPTIMIZATION.

Iowa State University, Ph.D., 1976
Operations Research

Computational aspects of linear multiple objective optimization

by

John Mark Trzeciak

A Dissertation Submitted to the

Graduate Faculty in Partial Fulfillment of

The Requirements for the Degree of

DOCTOR OF PHILOSOPHY

Department: Industrial Engineering
Major: Engineering Valuation

Approved:

Signature was redacted for privacy.

In Charge of Major Work

Signature was redacted for privacy.

For the Major Department

Signature was redacted for privacy.

For the Graduate College

Iowa State University
Ames, Iowa

1976

TABLE OF CONTENTS

1. INTRODUCTION

Consider the inherent deficiencies of "traditional" (i.e., single objective function) mathematical programming formulations of real-world models in view of the fact that a decision-maker functions in a multi-criterion environment. Virtually all decision-making situations involve simultaneous consideration of multiple and oftentimes conflicting "goals" or objectives.

Assuming that it is possible to construct mathematical expressions for a decision-maker's goals, the resulting formulation describes a multiple criteria programming problem. Clearly, multiple objective optimization models provide a superior representation of real-world decision-making situations relative to single objective models. Although the concept of multiple criterion optimization is intuitively appealing, the "solution" of multiple objective programming problems raises some serious questions with regard to theoretical and computational aspects of the problem. In particular, the criterion with which to judge optimality is itself subject to debate and controversy. Hence, any solution procedure used to identify optimal solutions must reflect this important theoretical consideration.

The primary emphasis of this study is focused on the computational aspects of multiple objective optimization problems involving linear functionals. An overview of this study is now presented.

1.1 Overview

Chapter 2 presents both a formal and an intuitive introduction to the general area of linear multiple objective optimization from a utility theoretic perspective. As a consequence of related work in decision theory, a linear multiple objective programming problem is recast as a (linear) vector maximum problem where the objective is the identification of solutions which are "admissable", "undominated", or "efficient". In particular, the computational aspects are considered in view of recent results in the literature. Also presented, in the spirit of background information, is the philosophically different approach to linear multiple objective optimization known as goal programming. Construction, analysis, and discussion of this problem focus on the computational aspects of the model and the fundamental issue of a "measure" of goal achievement is also addressed. Moreover, the equivalence of goal programming and linear regression (with or without side conditions) is established to provide motivation for studying alternative measures of goal achievement in view of the recent trend to consider alternative criterion of fit in linear regression.

The mathematical preliminaries and the intermediate results presented in Chapter 3 establish foundations for a new approach to the solution of the linear vector maximum problem based on the ℓ_2 metric. Although the solution procedure is developed in detail, the computational advantages of the approach are questionable. Other theoretical aspects of the approach are considered and discussed in detail.

The development and analysis presented in Chapter 3 provide a framework for the more general results contained in Chapter 4. In particular, a solution procedure is developed for the linear multiple objective problem based on the ℓ_p metric when $p \in [1,\infty)$. The analysis focuses initially on the linear vector maximum problem and is then extended to accommodate the more general goal programming problem. As a consequence of the ℓ_p metric, the solution procedure utilizes a branch of convex programming known as geometric programming. Motivation for the geometric programming formulation is derived from the computational advantages inherent in its associated dual problem. Moreover, it will be shown that the resulting dual problem can be solved by linear programming techniques.

Recognizing the importance of duality in mathematical programming problems, Chapter 5 provides a brief overview of an interesting dual problem associated with the a linearly constrained minimum norm problem. Based on results established by another author, a dual problem is considered which does not contribute to the computational aspects of the problem. However, analysis of this dual problem utilizing the Lagrangian function does provide a result which may be useful in interpreting the physical significance of the dual problem.

The last major chapter of the thesis, Chapter 6, summarizes the key results of the paper. In particular, the computational aspects of the dual problem presented in Chapter 4 are explored further as criticisms of this approach are considered.

1.2 Notation

To obtain notational consistency with related literature, the following conventions will hold throughout this thesis:

1. The symbol " $\Leftrightarrow$ " reads as "if and only if" or "is equivalent to."

2. Let $x,y \in R^n$. Then $x \geqq y \Leftrightarrow x_j \geqq y_j$, $j=1,\ldots,n$.

3. Let $x,y \in R^n$. Then $x \overset{>}{-} y \Leftrightarrow x \geqq y$, $x \neq y$.

4. Let $x,y \in R^n$. Then $x > y \Leftrightarrow x_j > y_j$, $j=1,\ldots,n$.

5. $\ell n(a) = \log_e(a)$ (for $a > 0$).

2. LINEAR MULTIPLE OBJECTIVE OPTIMIZATION

The roots of multiple objective optimization are found in the literature of classical physics, astronomy, and also in the related literature on the theory of games, decisions, and utility. In the context of a mathematical programming problem, multiple objective optimization is, in a broad sense, concerned with the constrained maximization of some measure of achievement or utility. To establish the relationship between the theory of utility and multiple objective optimization consider the following construction of a linear multiple objective programming problem.

Let

$$g_1(x),\ g_2(x),\ \ldots,\ g_k(x)$$

denote a set of linear real-valued objective functions. Here

$$g_i(x) = c_i x \qquad\qquad i=1,\ldots,k$$

where $c_i \varepsilon R^n$ represent vectors of known constants and $x \varepsilon R^n$ is a vector of unknown decision variables. Also, define

$$S = \{x \mid x \varepsilon R^n,\ Ax \leqq b,\ x \geqq 0\}.$$

Thus, S describes a set of linear side conditions, or constraints, on the model. Consider, also, the function

$$g(x) = \begin{bmatrix} g_1(x) \\ g_2(x) \\ \vdots \\ g_k(x) \end{bmatrix} = Cx$$

so that C is a $k \times n$ matrix of known constants. Note that the vector-valued function g provides a mapping $g: R^n \to R^k$ from the "decision space" to an "outcome space" as defined by the linear operator C. Utilizing this notation we now state the most general form of our linear multiple objective programming problem.

PROBLEM U

Maximize $U(Z)$

subject to

$$Cx = Z$$

$$x \in S.$$

Here the real-valued function U, $U: R^k \to R$, assigns a measure of utility for the decision-maker, given the outcome vector Z defined by a particular decision vector $x \varepsilon S$.

A complete treatment of the theory of utility is beyond the scope of this thesis; a more complete discussion of this subject may be found in Von Neumann and Mortgenstern [43], Hadley [25], or Chernoff and Moses [12]. However, it is instructive and relevant to briefly outline what is involved with the existence and construction of such a utility function. With regard to the existence of a utility function, consider the following four axioms as presented in Chernoff and Moses [12]:

1. With sufficient calculation an individual faced with two prospects P_1 and P_2 will be able to decide whether he likes each equally well, or whether he prefers P_2 to P_1.

2. If P_1 is regarded at least as well as P_2 and P_2 at least as well as P_3, then P_1 is regarded at least as well as P_3.

3. If P_1 is preferred to P_2 which is preferred to P_3, then there is a mixture of P_1 and P_3 which is preferred to P_2, and there is a mixture of P_1 and P_3 over which P_2 is preferred.

4. Suppose the individual prefers P_1 to P_2 and P_3 is another prospect. Then we assume that the individual will prefer a mixture of P_1 and P_3 to the same mixture of

P_2 and P_3.

Von Neumann and Mortgenstern [43] have shown that if a decision-maker can satisfy these four axioms then the decision-maker has a utility function $U(U:R^k \rightarrow R)$ which satisfies the following:

PROPERTY 2.1 If

1. $Z_1, Z_2 \varepsilon R^k$ and
2. $Z_1 \geq Z_2$

then

$$U(Z_1) > U(Z_2).$$

In view of Problem U, Property 2.1 can be interpreted as a criterion with which to measure optimality when one is concerned with the simultaneous maximization of the k linear objective functions. Note that the four axioms only address the existence of a utility function--the behavior of which is described in Property 2.1. The actual construction of a particular utility function is indeed a difficult and complex task. (For a discussion of the complexities associated with the construction of a utility function see Brandis [5].)

Although it is apparent that the fundamental nature of multiple objective optimization is embodied in the theory of utility, it is clear that a more pragmatic approach to the development of a measure of achievement is not only desirable but also a necessity. Moreover, it is evident that the acceptance of multiple objective optimization as a decision-making tool depends critically on the development of relatively straightforward solution techniques for use by the decision-maker.

Following, in spirit, a utility function approach, we now present two distinctly different approaches to linear multiple objective optimization. In particular, we will review the classic vector maximum problem and another problem of more recent vintage known as goal programming. In each case we will discuss the relative merits of the approach and note that these procedures, in fact, were developed to circumvent the complexities of multiple criterion decision-making. To provide motivation for considering alternative measures of achievement and to establish a foundation for the material presented in Chapter 4, we will also outline the classic statistical problem of constrained regression.

2.1 The Vector Maximum Problem

The vector maximum problem first appeared in the literature of mathematical programming in the classic paper by Kuhn and Tucker [33] on nonlinear programming. Recognizing the importance of multiple objective optimization, Kuhn and Tucker developed a set of necessary and sufficient optimality criteria for the following problem.

PROBLEM VM (Kuhn and Tucker [33]) To find an x^o that maximizes the vector function Gx constrained by $Fx \geqq 0$, $x \geqq 0$--that is, to find an x^o satisfying the constraints and such that $Gx \geq Gx^o$ for no x satisfying the constraints.

With regard to our formulation of the linear multiple objective programming problem, the vector function Gx corresponds to our g(x) and the constraints $Fx \geqq 0$, $x \geqq 0$ correspond to our set S. Thus, we focus attention on a linear version of Problem VM. It is important to note that the construction of a specific utility function, U, is avoided. Moreover, in view of Property 2.1, it is assumed that such a utility function exists and that the measure of utility is maximized when the outcome vector Z is "maximized". The linear version of the vector maximum problem may be expressed as:

PROBLEM LVM

"Maximize" Cx

subject to

$x \varepsilon S$.

Note that the function Cx is vector-valued and that the objective to "maximize" this function does not conform to a traditional criterion of optimality. Thus, to solve Problem LVM we need to recast the concept of optimality and establish a new criterion for identifying those vectors which are, in some sense, optimal. These "optimal" vectors are described in the following definition of efficiency.

DEFINITION A point $x^o \varepsilon S$ is said to be efficient if, and only if, there does not exist another $x \varepsilon S$ such that $Cx \geq Cx^o$.

An intuitive interpretation of an efficient point could be given as that point $x^o \varepsilon S$ which is undominated by all other points $x \varepsilon S$ to the extent that an increase in one of the components of $Z^o = Cx^o$, say Z_i^o, is made only at the expense of a decrease in at least one other component of Z^o, say Z_j^o. Thus, efficient solutions are analogous to "Pareto optimal solutions", "admissable points", and "proper solutions" in the context of economics, decision theory, and related areas as studied by Karlin [27], Von Neumann and Mortgenstern [43], Geoffrion [23], Kuhn and Tucker [33] and others. As will be shown in subsequent sections of this chapter, however, the solution of Problem LVM under the criterion of efficiency is, indeed, a significant computational task.

Before presenting some relevant results on a general solution of Problem LVM, it is instructive to consider the most common solution procedure illustrating the fact that it is designed to circumvent the true problem of multiple objective optimization. First, observe that Cx is a vector-valued function and that no generalized solution procedure is currently available to solve Problem LVM in the sense that the simplex algorithm is readily available for linear programming. Now, consider a cardinal ranking of the goals

$g_1, g_2, \ldots, g_k$ as determined by a weighting vector $v \epsilon R^k$ with components $v_i > 0$. This weighting of the objective functions suggests the following computationally attractive variant of Problem LVM:

PROBLEM LVMW

$$\text{Maximize } v'Cx$$

subject to

$$x \epsilon S.$$

Since $v \epsilon R^k$ is a vector of known constants, it follows that $v'C: R^n \to R$. Thus, Problem LVMW is amenable to the methods of linear programming since $v'Cx$ describes a linear real-valued function. Philip [39] has shown that this approach will identify an efficient solution to Problem LVM by establishing the following result.

LEMMA 2.2 A point $x^o \epsilon S$ is said to be efficient (for Problem LVM) if, and only if, there exists a vector $v \epsilon R^k$ such that

1. $v_i > 0, \qquad i=1,2,\ldots,k$
2. $\sum_{i=1}^{k} v_i = 1$
3. x^o solves maximize $\{v'Cx \mid x \epsilon S\}$.

Implicit in this procedure, however, is the critical assumption that the decision-maker has a prior knowledge of the relative merits of each of the k objectives and that these relative "weights" are accurately defined by the vector v. Clearly, this is a very strong condition which must be satisfied. Realistically, this assumption is often too strong to meet, but the decision-maker is sometimes forced to proceed with this weighting approach because of its computational advantages. Further analysis of the weighting vector approach, in view of Lemma 2.2, reveals that a decision-maker, in theory, could identify all efficient points to the problem given that one could generate all possible vectors $v \epsilon R^k$ which satisfy the conditions of the lemma. It is immediately obvious that generating the set of all possible weighting vectors is futile. Although this procedure is particularly attractive, in view of the Lemma 2.2, it affords little promise for the decision-maker who is not prepared to establish absolute rankings but is merely interested in identifying a set of "admissable" solutions.

Perhaps some of the most significant developments on the general theory of linear multiple objective programming can be attributed to Steuer [42] which later appeared in the literature by Evans and Steuer [19,20]. In pursuit of a general procedure to identify all efficient solutions to Problem LVM, Steuer developed an algorithmic approach based on the application of several well-known theorems of the alternative (see, for example, Mangasarian [37]). Motivated by the premise that a decision-maker would like to select his best "compromise" solution from among the set of all efficient solutions, the procedure attempts a characterization of the set:

$$E = \{x \mid x \epsilon S \text{ and } x \text{ is efficient}\}.$$

Since there could very well be an infinite number of points in the set E, the procedure was developed to identify all <u>basic</u> solutions of S which are efficient since this resulting subset E_B $(E_B \subseteq E)$ is guaranteed to have a finite number of elements. Moreover, it is proposed that E_B will provide a meaningful characterization of the set E. The algorithm is based on the construction of a subproblem at selected basic solutions to test not only for efficiency but also for an efficient direction in which to move to identify an adjacent basic solution which is also efficient. A brief description of this procedure is now presented.

<u>DEFINITION</u> A direction $\bar{\mu} \varepsilon R^n$ is a feasible direction at a point $x^* \varepsilon S$ if, and only if, there exists a scalar $\bar{\alpha} > 0$ such that $(x^* + \alpha\bar{\mu}) \varepsilon S$ for all $\alpha \varepsilon [0, \bar{\alpha}]$.

A logical extension of the concept of a feasible direction is given as follows.

<u>DEFINITION</u> A vector $\bar{\mu} \varepsilon R^n$ defines an efficient direction at a point $x^* \varepsilon S$ if, and only if,

1. $\bar{\mu}$ is a feasible direction at x^*, and
2. There exists a scalar $\alpha^* > 0$ such that $(x^* + \alpha\bar{\mu})$ is an efficient solution for all $\alpha \varepsilon [0, \alpha^*]$.

The relationship among efficient and feasible directions and efficient solutions is established in the following result.

LEMMA 2.3 (Evans and Steuer [20]) Let x^o be an efficient solution to Problem LVM and let $\bar{\mu}\varepsilon R^n$ be a feasible direction at x^o. Then $\bar{\mu}$ is an efficient direction at x^o if, and only if, there does not exist a feasible direction $\mu\varepsilon R^n$ such that $C\mu \geq C\bar{\mu}$.

Given that one has identified an efficient solution to Problem LVM, a subproblem must be constructed to determine the efficient direction $\bar{\mu}$. Recall that a basic solution is an extreme point of the convex polyhedron $S = \{x | x\varepsilon R^n, Ax \leq b, x \geq 0\}$. Suppose one has identified an efficient extreme point $x^o\varepsilon S$. Let A be partitioned into B (the basic column vectors of A) and N (the nonbasic columns of A). Likewise, let C be partitioned into C_B (the column vectors of C associated with the basic variables) and C_N (the column vectors of C associated with the nonbasic variables) and so on for

$$\mu = \begin{bmatrix} \mu_B \\ \mu_N \end{bmatrix}, \qquad \bar{\mu} = \begin{bmatrix} \bar{\mu}_B \\ \bar{\mu}_N \end{bmatrix}$$

This notation permits the following definition of the reduced cost matrix

$$W = C_B B^{-1} N - C_N$$

which is used to define an efficient direction as seen in the following result.

LEMMA 2.4 (Evans and Steuer [20]) Let x^o be an efficient solution to Problem LVM with associated basis matrix B. Then $\bar{\mu} \varepsilon R^n$ is an efficient direction at x^o if, and only if, there does not exist a feasible direction $\mu \varepsilon R^n$, at x^o, such that

$$W\mu_N \leq W\bar{\mu}_N .$$

Observe that feasible directions may be viewed as the edges of the polyhedron S adjacent to the point x^o. In determining the efficiency of such a feasible direction the procedure becomes somewhat complicated in the presence of degeneracy (see [19]) because in this case the number of extreme points adjacent to a given extreme point exceeds the number of nonbasic variables. To circumvent this problem a condition is enforced on $\mu \varepsilon R^n$, at a given extreme point x^o, which states that μ must satisfy

$$(-B^{-1}N)_D \, \mu_N \geqq 0$$
$$\mu_N \geqq 0$$

where $(-B^{-1}N)_D$ denotes the rows of $-B^{-1}N$ associated with the basic variables which are degenerate. Incorporating this condition into Lemma 2.4 results in the following test for efficiency of a direction μ.

LEMMA 2.5 (Evans and Steuer [20]) A vector $\bar{\mu} \epsilon R^n$ describes an efficient direction if, and only if, the system

$$W\mu_N \leq W\bar{\mu}_N$$
$$(-B^{-1}N)_D \, \mu_N \geqq 0$$
$$\mu_N \geqq 0$$

is inconsistent.

This stronger version of Lemma 2.4 incorporates a set of conditions relating to a pivoting strategy in the presence of degeneracy. To illustrate this point, consider the simplex tableau in a state of complete degeneracy. If the jth column vector, P_j, is chosen as the vector to enter the basis, then one may in fact "pivot" on any element of P_j which is nonzero without loss of feasibility. Such is the nature of the additional restrictions in Lemma 2.5

Focusing attention for the moment on the efficiency of a particular point x^o we have the following subproblem.

LEMMA 2.6 (Evans and Steuer [20]) A solution x^o is an efficient solution if, and only if, the problem

$$\text{maximize} \quad e'v$$

subject to

$$Wr + v = 0$$
$$(B^{-1}N)_D\, r + s = 0$$
$$r, v, s \geqq 0$$

is bounded where e' represents the sum vector of appropriate length. Now, addressing the issue of efficiency of a direction, we have the following subproblem test.

LEMMA 2.7 (Evans and Steuer [20]) Let x^o be an efficient extreme point. Then the subproblem given in Lemma 2.5 is consistent if, and only if, the subproblem

$$\text{maximize} \quad e'v$$

subject to

$$Wr + (-W\bar{\mu}_N)w + v = 0$$
$$(B^{-1}N)_D\, r + s = 0$$
$$r, w, v, s \geqq 0$$

has a feasible solution with $e'v > 0$. That is, the objective function is unbounded.

In this particular algorithm, each edge of the polyhedron S adjacent to an efficient point x^o is tested for efficiency. This is accomplished using the Chernikova procedure [11] to generate the set of all edges emanating from x^o where each edge can be viewed as a direction which is tested for efficiency using Lemma 2.7. From a computational perspective, the procedure is programmed into the following three phases:

1. Identify a basic feasible solution if one exists or terminate.

2. From a basic feasible solution, proceed to identify an efficient extreme point.

3. From an efficient extreme point, generate a list of all efficient extreme points.

Clearly, phase 3 is the most complex task. It is at this phase of the computation that subproblem construction and solution are determined. Furthermore, as discussed in [19], this procedure requires extensive bookkeeping if one is to ensure finiteness of the algorithm. Phases 1 and 2 deserve special note since the procedure developed supports five options with which to identify the initial efficient solution. Perhaps the most computationally attractive approach would be to assign some arbitrary weighting of the objectives and, in view of Lemma 2.2, solve a linear model. However, it is obvious that in using this method one runs the risk of overlooking a basic efficient solution in pursuit of that particular basic efficient solution which maximizes $v'Cx$.

Although theoretically sound, a weakness in the philosophy of the solution procedure is evident since the resulting characterization E_B of E may be quite large albeit finite. This situation is analogous to providing the decision-maker with too much information. It does, however, address the problem of multiple objective programming where the decision-maker is not in a position to establish an absolute ranking of the goals or objectives.

Other authors active in the area of efficient set methods include Markowitz [38], Geoffrion [23], and Karlin [27]. In particular, Geoffrion [23] proposed a procedure to identify all efficient solutions to a bi-criterion program (two objective functions). However, the results of his work have not been extended to problems of a more general nature.

A popular variant of the pure linear vector maximum problem has come to be known as goal programming. Here, a decision-maker's measure of utility is maximized when a measure of "goal achievement" is maximized. We now present some relevant results in this area.

2.2 Goal Programming

A philosophically different approach to linear multiple objective programming was proposed by Charnes and Cooper [7]. Not only does this approach afford significant computational advantages, but it also provides a more realistic model of many real-world decision-making situations. Consider, again, the most general form of our linear multiple objective programming problem.

PROBLEM U

$$\text{maximize} \quad U(Z)$$

subject to

$$Cx = Z$$

$$x \in S.$$

A fundamental assumption implicit in goal programming is that the utility function, $U(Z)$, is maximized when the outcome vector $Z \in R^k$ gets as close as possible to some target or "goal" vector $g^* \in R^k$ which is assumed to be known and constant. Thus, the utility function is never explicitly constructed but is assumed to exist and, by definition, it provides a measure of "goal attainment".

As introduced by Charnes and Cooper [7], discussed by Charnes and Cooper [6,7], and applied by Charnes, Cooper, Klingman and Niehaus [8] and Charnes, Cooper, Niehaus and Scholtz [9], the measure of achievement is maximized when the distance between the point $g^* \epsilon R^k$ and the point $g(x^o) \epsilon R^k$, $g(x) = Cx$, is minimized. Recognize immediately that distance can be defined any number of ways. Chapters 3 and 4 exploit alternative measures of achievement by considering a more general definition of distance. Of immediate interest is the concept of distance between two vectors $x, y \epsilon R^k$ as defined by the metric

$$\ell_p: \left[\sum_{i=1}^{k} |x_i - y_i|^p \right]^{1/p}$$

when $p = 1$. Thus, in goal programming the measure of achievement is maximized when the metric

$$\sum_{i=1}^{k} |g_i(x) - g_i^*|$$

is minimized where $(g_i(x), g_i^*)$ denote a goal function and its associated target value or goal. Under the assumption that this ℓ_1 metric accurately describes the decision-maker's utility, the following linear programming model can be employed.

PROBLEM GP1

$$\text{minimize} \quad \sum_{i=1}^{k} (w_i^+ d_i^+ + w_i^- d_i^-)$$

subject to

$$Cx + d^- - d^+ = g^*$$

$$x \varepsilon S$$

$$d^+, d^- \varepsilon R^k, \; d^+, d^- \geqq 0$$

$$\langle d^-, d^+ \rangle = 0$$

The inclusion of weights (w_i^+, w_i^-) in the model is intended to provide the decision-maker the option of specifying the relative importance of the various goals. Without loss of generality we can assume that $w_i^+ = w_i^- = 1$ for all i and provide the following interpretation of the goal programming model. Consider any of the k linear goal function, say $g_i(x)$, and its corresponding goal g_i^*. In view of Figure 1, the nature of goal programming is to find that particular $x^o \varepsilon S$ that minimizes the sum of the deviations which describe the distance between $g_i(x)$ and g_i^*.

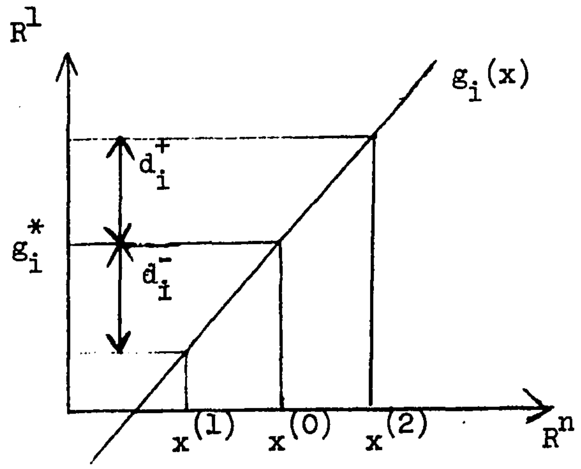

Figure 1. The Nature of Goal Programming

The last constraint in Problem GP1 ensures that a given goal cannot have both positive (d_i^+) and negative (d_i^-) deviations active simultaneously. Fortunately, this (nonlinear) condition will always be satisfied when the simplex algorithm is used to solve the model. (The definition of a basis excludes the possibility of two active linearly independent vectors.)

Inherent in the construction of Problem GP1 is the assumption that the weights (w_i^+, w_i^-) define a cardinal ranking of the goals. Moreover, it is also assumed that minimizing the resulting "weighted" ℓ_1 metric is equivalent to maximizing the decision-maker's measure of utility. For the sake of completeness it should be noted that an ordinal ranking of the goals is sometimes useful (see, for example, Lee [34] or Ijiri [26]). In this situation the goals are ranked according to some priority structure. That is, assume that goal g_i is ranked ordinally above another goal g_j. Then a priority level P_i is assigned to the deviational variables corresponding to g_i and a priority level P_j is assigned to the deviational variables associated with g_j such that $P_i >> P_j$. In this case as well, a minor variant of the simplex algorithm can be employed to solve the problem (see Lee [34]).

The model outlined above describes the tool which is used to solve virtually all goal programming problems encountered in practice. Computer code implenting this procedure has been developed by Lee and Hoffman in [34] and is generally well accepted. However, the practitioner interested in goal programming should address the following fundamental (and often overlooked) issue: "Does minimizing a weighted ℓ_1 metric accurately describe a meaningful measure of achievement?" Clearly, this approach to goal programming affords the luxury of a well-known and readily available solution procedure. However, one wonders if the acceptance of this technique is based on its merit as a model of a decision-maker's utility or whether its popularity is derived from the inherent linearity of the model. Resolution of this philsophical issue is beyond the scope and not the intent of this thesis. However, to provide further insight we will consider another goal programming model based on a different measure of achievement. The motivation behind introducing an alternative model is to illustrate the mathematical complexities one encounters with alternative measures of achievement and to provide a foundation for the results of Chapter 3.

A goal programming model based on a measure of achievement different from the ℓ_1 metric was proposed by Ijiri [26]. To illustrate the nature of his approach consider, again, Problem U where the utility function, U, is maximized when the metric

$$\ell_p: \left[\sum_{i=1}^{k} |x_i - y_i|^p \right]^{1/p}$$

is minimized for $p = 2$. Here the ℓ_2 metric defines the Euclidean distance between the two vectors $x, y \epsilon R^k$. More specifically, let

$$g(x) = \begin{bmatrix} g_1(x) \\ g_2(x) \\ \vdots \\ g_k(x) \end{bmatrix} = Cx \quad .$$

represent a set of linear goal functions defined as a vector-valued function g and let

$$g^* = \begin{bmatrix} g_1^* \\ g_2^* \\ \vdots \\ g_k^* \end{bmatrix}$$

represent a set of corresponding goals. Note that the ℓ_2 metric induces a Euclidean norm which we will denote as $||\cdot||_2$ (a more complete discussion of norms is given in Chapter 3). Thus, a general formulation of a goal programming model based on the ℓ_2 metric is given as:

PROBLEM GP2

$$\text{minimize} \quad ||g(x) - g^*||_2$$

subject to

$$x \in S.$$

Observe that Problem GP2 is inherently nonlinear and, in particular, is amenable to the methods of quadratic programming. In an effort to avoid the complexities associated with nonlinear optimization, Ijiri [26] proposes a generalized inverse approach to the solution of this model. To illustrate the nature of this approach let us assume, for the moment, that the program is unconstrained so that our problem simplifies to

$$\text{minimize} \quad \|Cx - g^*\|_2 \, ,$$

or equivalently

$$\text{minimize} \quad \left[\sum_{i=1}^{k} |g_i(x) - g_i^*|^2 \right]^{1/2} .$$

Consider the kxn matrix **C** and let C^* denote its <u>generalized inverse</u>. It then follows, as a consequence of the theory of a generalized inverse of an arbitrary matrix (see Ijiri [26]), that the vector

$$\hat{x} = C^* g^*$$

is that unique vector for which

1. $C\hat{x}$ is a vector which is the minimum Euclidean distance to g^* from among all vectors in R(C) (R(C) denotes row space of the matrix C), and

2. $\hat{x}$ is minimum Euclidean distance to the origin.

Hence, $\hat{x}$ solves the problem

$$\text{minimize} \quad ||x||_2$$

subject to

$$||Cx - g^*||_2 \text{ to be a minimum}$$

where $||\cdot||_2$ denotes the Euclidean norm. In Chapter 3 a more complete discussion of the properties of a generalized inverse is presented. However, it is instructive at this point to note that alternative solutions to Problem GP2 are obtainable. In this situation, we may dispose of the uniqueness property guaranteed by

$$\text{minimize} \quad ||x||_2$$

and consider the solution defined by the expression

$$x_\gamma = C^* g^* + (I - C^* C)\gamma, \quad \gamma \varepsilon R^n.$$

Here, x_γ solves minimize $||Cx - g^*||_2$ for any vector $\gamma \varepsilon R^n$. As will be shown in Chapter 3, $(I - C^* C)\gamma$ is a vector from the null space of the matrix C so that

$$\begin{aligned} Cx_\gamma &= C[C^* g^* + (I - C^* C)\gamma] \\ &= CC^* g^* + C(I - C^* C)\gamma \end{aligned}$$

$$= CC^* g^* + C\Upsilon - CC^* C\Upsilon$$

$$= CC^* g^* + C\Upsilon - C\Upsilon$$

$$= CC^* g^* = C\hat{x}$$

It is also important to note that these results require further extensions if one is to incorporate constraints as given in the statement of Problem GP2. Moreover, the construction of the generalized inverse of an arbitrary kxn matrix is itself a significant computational task. Although Ijiri's approach does allow for a measure of achievement different than the ℓ_1 metric, it does not afford the computational advantages of linear programming.

2.3 Goal Programming and Constrained Regression

As recently noted by Charnes and Cooper [6], the concept of goal programming or goal achievement is not totally new. A strongly equivalent problem studied by statisticians (see [1], [2], [44]) is that of constrained ℓ_p approximation using linear approximating functions. To show that ℓ_p approximation is completely equivalent, in spirit, to goal programming, consider the following statistical problem:

Let X denote a kx(n+1) "observation" matrix of known constants (k observations of n independent variables), let Y represent a kx1 vector of observations of some dependent variable, and let

$$\beta' = (\beta_o, \beta_1, \ldots, \beta_n)$$

denote a vector of unknown parameters to be estimated. Optionally, let

$$S = \{\beta \mid \beta \varepsilon R^{n+1},\ D\beta \leqq d\}$$

represent a set of linear side conditions on the vector β. The nature of regression may now be described as follows:

PROBLEM GR1 Find a vector β^o, if it exists, such that

1. $\beta^o \varepsilon S$ (if appropriate)

2. β^o maximizes some measure or criterion of fit for the model.

If one assumes that maximizing the measure or criterion of fit corresponds to minimizing the ℓ_p norm

$$||X\beta - Y||_p$$

then we can state the ℓp approximation problem as:

PROBLEM GR2

$$\text{minimize} \quad ||X\beta - Y||_p$$

subject to

1. $D\beta \leqq d$ (optional)

2. $\beta \epsilon R^{n+1}$

The relationship between goal programming and ℓ_p approximation is now evident if one considers the vector-valued function $X\beta$ to be a goal function and the observation vector Y to be the target or goal vector.

Perhaps the most frequently encountered regression problem is known as "least squares" regression. In view of Problem GR2, a least squares regression is defined as that ℓ_p approximation problem for which $p = 2$. Thus, in this context, least squares regression is then "regression under ℓ_2" which may be expressed as

$$\underset{\beta \epsilon R^{n+1}}{\text{minimize}} \quad ||X\beta - Y||_2$$

subject to

$$D\beta \leqq d \qquad \text{(optional)}$$

Most least squares regression problems do not involve the optional side conditions since this option adds significantly to the computational aspects of the problem as will be demonstrated shortly. Without side conditions, least squares regression reduces to finding a unique solution to the normal equations (see [15]) given as

$$X'X\beta = X'Y$$

which is provided by the computation

$$\beta^o = (X'X)^{-1} X'Y$$

It can be easily shown that β^o solves the problem

$$\text{minimize} \quad ||\beta||_2$$

subject to

$$||X\beta - Y||_2 \text{ to be a minimum}$$

which is completely equivalent to the goal programming model proposal by Ijiri where the measure of achievement is based on the ℓ_2 metric. Just as in Ijiri's model, the solution procedure for constrained regression (i.e., including side conditions on β) results in a nonlinear optimization problem which is amenable to the methods of quadratic programming.

It is becoming increasingly evident that the popular least squares (ℓ_2) regression technique does not always yield a superior estimate of the unknown parameter β. In particular, the ℓ_2 estimate of β is quite sensitive to outliers in the observed data. Moreover, it can also be shown that when the error terms associated with the observations do not follow the $N(0,1)$ distribution then some of the desirable properties of the ℓ_2 estimate are not satisfied (such as consistency, unbiasedness, maximum likelihood, etc.). Furthermore, if the error terms associated with the observations do not follow a multivariate normal distribution (with mean 0 and variance 1) then the distribution of the resulting ℓ_2 estimate may not follow a multivariate normal in which case the analysis of variance, as we know it today, is invalid. Recognizing the sensitivity of ℓ_2 estimates to outliers in the data, research in [1] and [22] is focused on the more general problem of ℓ_p regression. Although a key aspect of this research is the determination of the distributions of the estimates under certain assumptions regarding the distribution of the errors, the general ℓ_p regression problem is inherently a nonlinear optimization problem for virtually all values of p. Leaving the problem of distributions of estimates to statisticians, we will focus on the optimization problem used to identify a particular estimate of β under the criterion ℓ_p.

2.3.1 Regression Under ℓ_1 (LAVE):

The regression problem defined under the criterion ℓ_p when $p = 1$ is known as Least Absolute Value Estimation which may be expressed as follows:

$$\underset{\beta \in R^{n+1}}{\text{minimize}} \quad ||X\beta - Y||_{p=1}$$

subject to

$$D\beta \leqq d \qquad \text{(optional)}$$

Note that this formulation is completely equivalent to goal programming based on the ℓ_1 metric. Hence, LAVE is amenable to the methods of linear programming.

2.3.2 Regression Under ℓ_∞ (MINIMAX):

Sometimes one is interested in the estimate of β where the criterion of fit is given as the ℓ_p metric when $p = \infty$. It can be shown that this is equivalent to

$$\text{minimize} \quad \{\sup_i |(X\beta)_i - Y_i)|\}$$

which is known as the Chebychev criterion and describes a minimization of the absolute value of the maximum deviation. In this case, taking the limit as $p\to\infty$ in the formulation given as Problem GR2 results in the computationally equivalent problem:

$$\text{minimize} \quad e$$

subject to

$$D\beta \leq d \qquad \text{(optional)}$$
$$-e \leq (X\beta)_i - Y_i \leq e \qquad \text{for all } i$$
$$\beta\varepsilon R^{n+1},\ e\varepsilon R,\ e \geq 0.$$

which is, clearly, amenable to the methods of linear programming.

2.3.3 Regression Under ℓ_p:

Recognize that regression under ℓ_1 and regression under ℓ_∞

are just particular problems from the more general family of regression problems based on the ℓ_p metric. Typically, in ℓ_p approximation one is concerned with regression problems based on the ℓ_p metric when $p \in [1,\infty]$. For $p \in (0,1)$ the ℓ_p metric does not induce a norm because the triangle inequality is reversed (see Section 3.1). Moreover, for $p \in (0,1)$ the resulting optimization problem is not within the domain of convex programming because the objective function to be minimized is concave in the parameter β. In general, we will restrict attention to the case where $p \geq 1$--the computational aspects of which are summarized as follows:

p	Estimate	Solution Procedure
1	LAVE	linear programming
$1 < p < \infty$	ℓ_p	nonlinear (convex) programming
∞	MINIMAX	linear programming

It follows that the most common regression problems are based on the ℓ_p metric where $p = 2 \pm \varepsilon$. That is, it is often instructive to look at the resulting estimates which are "almost ℓ_2" in the sense that p is specified to be in some ε-neighborhood of 2. In particular, empirical results by Forsythe [22] suggest a strong case for ℓ_p approximation where $p = 1.5$ under certain conditions concerning outliers in the data. Unfortunately, his resulting model is not amenable to the methods of linear programming.

In this study, Forsythe employed the gradient projection method of Fletcher and Powell (see [22]). Although this optimization procedure is theoretically sound, it requires the evaluation of a derived function (first order derivative) and other complexities associated with nonlinear programming.

We will not pursue the subject of ℓ_p approximation further since the intent of this analysis is to demonstrate the strong equivalence between discrete ℓ_p approximation and what may be termed "generalized" goal programming. Just as regression has historically been based on the ℓ_2 metric (without side conditions) so has goal programming been based, for the most part, on the ℓ_1 metric. In both cases, these approaches afford significant computational advantages. Namely, the resulting problems are inherently linear.

3. A MINIMUM NORM APPROACH TO VECTOR MAXIMIZATION

In this chapter we will revisit the linear version of the vector maximum problem and develop a solution procedure to identify efficient solutions. The results contained herein do not contribute significantly to the computational aspects of the problem. The primary emphasis of this chapter is the development of new insights with regard to the complex nature of multiple criterion decision-making. Moreover, these results are used as a foundation for the material presented in Chapter 4 where stronger computational results are presented for a more general problem.

3.1 Mathematical Preliminaries

In view of the fact that the results of this chapter are based on minimizing the norm of a vector (hence, a minimum norm approach) we now present the following well-known results on norms and general solutions to linear systems.

DEFINITION 3.1 Let X be a linear vector space. Then a real-valued function, denoted by $||.||$, which maps each element x in X into a real number is called the norm of x if it satisfies the following axioms:

1. $||x|| \geq 0$ for all $x \epsilon X$,
2. $||x|| = 0$ if, and only if, $x = 0$
3. $||\alpha x|| = |\alpha| . ||x||$ for all $\alpha \epsilon R$ and each $x \epsilon X$, and

4. $||x + y|| \leq ||x|| + ||y||$ for each $x,y \in X$ (triangle equality).

Clearly, the norm is an abstraction of our usual concept of length. In particular, $||x||$ defines a measure of length from the point x, in some vector space, X, to the origin. Likewise, $||x-y||$ defines a measure of length between the two points x and y in some vector space. Note that there exists a spectrum of functions which satisfy the properties of a norm. For purposes of our discussion, we will utilize the well-known metric

$$\ell_p : \left[\Sigma_i |x_i|^p \right]^{1/p}$$

which can be shown to induce a norm for $p \geq 1$. We will also have need of the following property of norms.

PROPERTY 3.2 Let x and y be any two elements of a normed linear vector space. Then

$$||x|| - ||y|| \leq ||x - y||.$$

The solution procedure to be presented is based on a minimum norm problem where the particular norm of interest defines the Euclidean distance. Furthermore, the approach is based on the generalized inverse approach to the solution of a linear system. A conceptual interpretation of the relationship between Euclidean distance and the generalized inverse solution to a linear system is now presented.

Let the linear operator C define a mapping $C: R^n \to R^k$ so that C is a $k x n$ matrix and consider the two vectors

$$Z^*, Cx^* \epsilon R^k$$

where $x^* \epsilon R^n$. If there exists an $x^* \epsilon R^n$ such that

$$Cx^* = Z^*$$

then Z^* is said to be an element of the row space of the matrix C. We represent this situation notationally as $Z^* \epsilon R(C)$ and note that

$$\underset{x \epsilon R^n}{\text{minimum}} \; ||Cx - Z^*|| = 0 \quad \text{when} \quad Z^* \epsilon R(C)$$

as a consequence of Definition 3.1. However, if $Z^* \notin R(C)$ then there does not exist a vector $x \epsilon R^n$ such that

$$Cx = Z^*.$$

In this case,

$$\underset{x \epsilon R^n}{\text{minimum}} \; ||Cx - Z^*|| > 0 \quad \text{when} \quad Z^* \notin R(C).$$

Now, let $\hat{Z} \varepsilon R(C)$ be arbitrary and consider the general linear system

$$\hat{Z} = Cx. \tag{3.1}$$

Then a solution, $\hat{x}$, to (3.1) is given as

$$\hat{x} = C^{*}\hat{Z},$$

where C^{*} denotes the nxk generalized inverse of the kxn matrix C, since

$$C\hat{x} = CC^{*}\hat{Z} = \hat{Z}$$

by the inverting property of the generalized inverse (see Appendix A). This solution $\hat{x}$ is, in general, not the only solution to the system (3.1). One possible approach to the identification of all possible solutions is based on the concept of a null space of the transformation (i.e., matrix) C.

<u>DEFINITION 3.3</u> Let C be an arbitrary kxn matrix. Then a vector $x^{o} \varepsilon R^{n}$ is said to be an element from the null space of the matrix C (i.e., $x^{o} \varepsilon N(C)$) if, and only if,

$$Cx^{o} = 0.$$

With regard to the solution of system (3.1), if $\hat{x}$ is a solution to

$$Cx = \hat{Z}$$

then any vector

$$x = \hat{x} + x^{o}$$

is also a solution to system (3.1) if, and only if, $x^{o} \epsilon N(C)$ since

$$Cx = C(\hat{x} + x^{o}) = C\hat{x} + Cx^{o} = C\hat{x} = \hat{Z}$$

Thus, the set of all solutions to (3.1) can be obtained by adding each vector in $N(C)$ to $\hat{x}$. The uniqueness of a solution to (3.1) depends entirely upon whether or not $N(C)$ consists of only the null vector (i.e., $x = 0$) which is true if, and only if, C is nonsingular. In pursuit of a procedure to identify these alternative solutions we present, without verification, the following intermediate result. Namely,

$$\{x | x \epsilon R^{n}, \hat{Z} = Cx\} = \{x | x \epsilon R^{n}, x = C^{*}\hat{Z} + (I - C^{*}C)\gamma, \gamma \epsilon R^{n}\}. \quad (3.2)$$

Here, $(I - C^{*}C)\gamma$ defines a vector from the null space of the matrix C (see [26]). Moreover, by allowing γ to span R^{n} we can obtain every vector in $N(C)$ and hence, in view of (3.2), all possible solutions to (3.1).

It is often the case, however, that we are interested in the linear system

$$Z = Cx$$

where $Z \notin R(C)$. This, of course, implies that there does not exist a vector $x \in R^n$ such that $Z = Cx$ (i.e., the system is inconsistent). The "least squares" property of the generalized inverse C^* now becomes an important issue. Assume $Z \in R^k$ is an arbitrary vector such that $Z \notin R(C)$ and consider the transformation

$$Z^* = CC^*Z.$$

It is a consequence of the generalized inverse that Z^* is an element of $R(C)$. Furthermore, Z^* is that unique vector which has minimum Euclidean distance to Z from among <u>all</u> vectors in $R(C)$. Thus,

$$\hat{x} = C^*Z$$

is a solution to the linear system

$$Z^* = Cx.$$

This least squares property of generalized inverses can, perhaps, best be illustrated by considering an equivalent optimization problem. Since this property will be useful later in the development, we present the following results for further reference.

<u>LEMMA 3.4</u> Let C be any $k \times n$ matrix and let $Z \varepsilon R^k$ be an arbitrary vector. Then there exists a vector $\hat{x}$, $\hat{x} \varepsilon R^n$, given as

$$\hat{x} = C^* Z$$

which is a (unique) solution to

$$\text{minimize} \quad ||x||$$

subject to

$$||Cx - Z|| \quad \text{to be a minimum}$$

where $||.||$ denotes the Euclidean norm.

<u>COROLLARY 3.5</u> Let $\hat{x} = C^* Z$ so that $||C\hat{x} - Z|| = \alpha$. Then

$$\{x \mid ||Cx - Z|| = \alpha\} = \{x \mid x = C^* Z + (I - C^* C)\gamma, \gamma \varepsilon R^n\}.$$

We will now proceed to construct and analyze a minimum norm problem to identify a class of solutions to a linear version of the vector maximum problem. Particular attention will be focused on alternative optimal solutions to the problem in view of Corollary 3.5.

3.2 A Classification of Efficiency

The procedures developed by Evans and Steuer [20] are based on the characterization of the set E (i.e., the set of all efficient solutions) by identifying the elements of the set E_B(E_B = $\{x \mid x \varepsilon E$ and x an extreme point of $S\}$, $S = \{x \mid x \varepsilon R^n, Ax \leqq b, x \geqq 0\}$). As an alternative we will develop a procedure to characterize the set E by appealing to a subset $E_{\ell 2}$ defined below. The results of this chapter assume that the following condition is satisfied for the vector maximum problem, Problem LVM.

CONDITION 3.6 Assume there exists a vector $Z^o \varepsilon R^k$ such that $Z^o > Cx$ for all $x \varepsilon S = \{x \mid Ax \leqq b, x \geqq 0\}$.

This condition will be given further attention later in this chapter. Moreover, Chapter 4 provides the results needed to relax this assumption.

In view of fact that Euclidean distance is determined by the ℓ_p metric for $p = 2$, consider the following extension of the concept of efficiency.

DEFINITION 3.7 A point $x^o \varepsilon S$ is said to be ℓ_2-efficient if, and only if, there does not exist another $x \varepsilon S$ such that

$$||Cx - Z^o|| < ||Cx^o - Z^o||$$

when Z^o is a vector which satisfies Condition 3.6 and $||.||$ denotes the Euclidean norm.

Just as E_B relates to and characterizes E we now choose to characterize E by appealing to a subset $E_{\ell 2}$ defined as

$$E_{\ell 2} = \{x | x \varepsilon S \text{ and } x \text{ is } \ell_2\text{-efficient}\}.$$

Establishing the relationship between $E_{\ell 2}$ and E we have the following result.

<u>LEMMA 3.8</u> $E_{\ell 2} \subseteq E$.

<u>Proof</u> Let $x^o \varepsilon E_{\ell 2}$ be arbitrary and assume $x^o \notin E$. It then follows by definition that there exists some point $x^* \varepsilon S$ such that

$$Z^o > Cx^* \geq Cx^o.$$

Hence,

$$||Cx^* - Z^o|| < ||Cx^o - Z^o||.$$

But this implies that $x^o \notin E_{\ell 2}$ and the result follows.

We begin the development with a restatement of a linear version of the vector maximum problem given as:

PROBLEM A

$$\text{"maximize"} \quad Cx$$

subject to

$$x\varepsilon S = \{x | x\varepsilon R^n, \ Ax \leq b, \ x \geq 0\}.$$

The definition of ℓ_2-efficiency suggests the following minimum norm problem.

PROBLEM B Find an x^o, if it exists, for which

$$||Cx^o - Z^o|| = \underset{x\varepsilon S}{\text{minimum}} \ ||Cx - Z^o||$$

The relationship between the minimum norm problem, Problem B, and the vector maximum problem, Problem A, is established by the following result.

LEMMA 3.9 If there exists a solution x^o which solves Problem B then

1. x^o is ℓ_2-efficient for Problem A, and

2. x^o is efficient for Problem A.

Proof The result follows immediately from Definition 3.7 and Lemma 3.8.

Recall from Section 3.1 that Cx_γ, where x_γ is given as

$$x_\gamma = C^* Z^o + (I - C^* C)\gamma, \gamma \varepsilon R^n,$$

defines that vector in $R(C)$ which is a minimum Euclidean distance to Z^o. In view of this result, consider the following problem and corresponding Lemma.

PROBLEM C Find a γ, if it exists, such that

1. $Ax_\gamma = A(C^* Z^o + (I - C^* C)\gamma) \leqq b$,

2. $x_\gamma = C^* Z^o + (I - C^* C)\gamma \geqq 0$,

3. $\gamma \varepsilon R^n$.

LEMMA 3.10 If there exists a solution γ^o which solves Problem C then

1. x_{γ^o} solves Problem B, and

2. x_{γ^o} is ℓ_2-efficient for Problem A.

Proof. If γ^o solves Problem C then, clearly, $x_{\gamma^o} \varepsilon S$ so x_{γ^o} is feasible for Problem B. Moreover, in view of Corollary 3.5, x_{γ^o} is also optimal for Problem B and hence, by Lemma 3.9, x_{γ^o} is ℓ_2-efficient for Problem A.

Although Problem C is amenable to the methods of linear programming, this intermediate result is not sufficient. In particular, it is quite possible that there may not exist a feasible solution to Problem C even though there exists solutions to Problem B. Infeasibility of Problem C indicates that

$$\{x_\gamma \,|\, Cx_\gamma \varepsilon R(C), ||Cx_\gamma - Z^o|| \text{ is minimum}\} \cap S = \emptyset.$$

Consider the following variant of Problem C.

PROBLEM D

minimize $||\delta||$

subject to

1. $||Cx - (Z^o - \delta)|| = 0,$ (3.5)

2. $x \varepsilon S,$ (3.6)

3. $\delta \varepsilon R^k, \quad \delta \geq 0.$ (3.7)

With regard to the feasibility of Problem D, we have the following result.

LEMMA 3.11 Assume Condition 3.6 is satisfied. Then Problem D is feasible if, and only if, Problem A is feasible.

Proof. Let (x^*, δ^*) denote any feasible solution to Problem D. Then, in view of (3.6), $x^* \varepsilon S$ and hence x^* is feasible for Problem A. Conversely, let x^* denote any feasible solution to Problem A. Then (3.6) holds and letting

$$\delta^* = Z^o - Cx^* \geqq 0$$

satisfies (3.5) and (3.7) under Condition 3.6.

With regard to the optimality of Problem D, we have the following intermediate result.

LEMMA 3.11 If (x^o, δ^o) solves Problem D and $\delta^o = 0$, then

1. x^o solves Problem B, and

2. x^o is ℓ_2-efficient for Problem A.

Proof. When $\delta^o = 0$, Problem D is equivalent to Problem C. Hence, the result follows as a consequence of Lemma 3.10.

Problem D has the following intuitive interpretation. If we can find a vector $Z^* \varepsilon R^k$, $Z^* = (Z^o - \delta^o) \varepsilon \Omega = \{Z | Z = Cx, x \varepsilon S\}$ which has minimum Euclidean distance to Z^o from among all vectors in Ω, then any $x \varepsilon S$ for which $||Cx - Z^*|| = 0$ solves Problem B. This problem can be simplified into a more computationally attractive problem by noting that

$$||Cx - (Z^{o} - \delta)|| = 0 \iff (Z^{o} - \delta)\varepsilon R(C).$$

But we can guarantee that $(Z^{o} - \delta)\varepsilon R(C)$ by enforcing the (linear) condition that

$$(Z^{o} - \delta) = CC^{*}(Z^{o} - \delta).$$

Furthermore, we can now exploit the generalized inverse by utilizing the fact that any x for which $||Cx - (Z^{o} - \delta)||$ is a minimum can be represented as

$$x_{\gamma,\delta} = C^{*}(Z^{o} - \delta) + (I - C^{*}C)\gamma$$

which suggests the following extension of Problem D.

PROBLEM E

minimize $||\delta||$

subject to

1. $Ax_{\gamma,\delta} = A[C^{*}(Z^{o} - \delta) + (I - C^{*}C)\gamma] \leqq b$
2. $x_{\gamma,\delta} = C^{*}(Z^{o} - \delta) + (I - C^{*}C)\gamma \geqq 0$

3. $(Z^{o} - \delta) = CC^{*}(Z^{o} - \delta)$

4. $\gamma \epsilon R^{n}$, $\delta \epsilon R^{k}$, $\delta \geqq 0$.

Since the particular norm employed in the objective function is the Euclidean norm and since δ must satisfy the restriction $\delta \geqq 0$, Problem E is amenable to the methods of quadratic (convex) programming. Moreover, available software will identify the optimal solution since the objective function assumes the quadratic form

$$\text{minimize} \quad y'Py + \underline{0}'y$$

where $y \epsilon R^{n+k}$,

$$y = \begin{bmatrix} \delta \\ \gamma \end{bmatrix}$$

and P is of the form

$$P = \left[\begin{array}{c|c} I_{nxn} & 0_{nxk} \\ \hline 0_{kxn} & 0_{kxk} \end{array}\right]$$

which is positive semi-definite. With regard to the optimality of Problem E, we have the following key result.

<u>LEMMA 3.12</u> If (γ^o,δ^o) solves Problem E, then

$$x_{\gamma^o,\delta^o} = C^*(Z^o - \delta^o) + (I - C^*C)\gamma^o$$

is ℓ_2-efficient for Problem A.

<u>Proof.</u> Note that if x^o solves Problem B then, in view of Lemma 3.9, x^o is ℓ_2-efficient for Problem A. Furthermore, since Problem E is equivalent to Problem D, it suffices to show that if (x^o,δ^o) solves Problem D then x^o solves Problem B. Thus, assume (x^o,δ^o) solves Problem D. Then

$$||Cx^o - (Z^o - \delta^o)|| = 0$$

if, and only if,

$$C^ix^o - Z_i^o + \delta_i^o = 0 \qquad \text{for} \qquad i=1,\ldots,k$$

or, equivalently

$$-C^ix^o + Z_i^o = \delta_i^o \qquad \text{for} \qquad i=1,\ldots k.$$

Thus, in matrix notation

$$-Cx^o + Z^o = \delta^o$$

or

$$||-Cx^o + Z^o|| = ||Cx^o - Z^o|| = ||\delta^o||.$$

But this implies that

$$\min_{\delta \geq 0} ||\delta|| = \min_{\delta \geq 0} ||Z^o - (Z^o - \delta)|| = \min_{x \in S} ||Cx - Z^o|| = ||Cx_{\gamma^o,\delta^o} - Z^o||.$$

Hence, $x^o_{\gamma^o,\delta^o}$ solves Problem B.

3.3 Characterization of $E_{\ell 2}$

Given that Problem E can be used to define a solution x^o, $x^o \epsilon E_{\ell 2}$, we would now like to characterize the set $E_{\ell 2}$. Let Ω denote the set of all feasible solutions to Problem E and let $(\gamma^o, \delta^o) \epsilon \Omega$ be optimal for Problem E. As a consequence of Corollary 3.5, for each $\gamma \epsilon R^n$ such that $(\gamma, \delta^o) \epsilon \Omega$ we have

$$x_\gamma = C^*(Z^o - \delta^o) + (I - C^*C)\gamma \qquad (3.8)$$

which defines an ℓ_2-efficient solution to Problem A. Moreover, in theory we can span the set $E_{\ell 2}$ by finding all possible $\gamma \epsilon R^n$ such that $(\gamma, \delta^o) \epsilon \Omega$ and x_γ (defined 3.8) is contained in $E_{\ell 2}$. Of course it may be possible that the subset $E_{\ell 2}$, $E_{\ell 2} \subseteq E$, contains an infinite number of points. Although we can, in theory, identify each element of $E_{\ell 2}$, we seek a procedure to guarantee finiteness of the algorithm. In view of this, it is proposed that the iteration of all alternative optimal solutions to Problem E will provide a meaningful characterization of the set $E_{\ell 2}$, the elements of which are determined by (3.8).

The approach presented in this chapter poses some serious computational questions. In particular, the formulation given as Problem E is not completely equivalent to Problem A in the sense that feasibility (and hence optimality) in Problem E does not guarantee feasibility (and hence optimality) in Problem A. Problem E is appropriate and valid under the assumption that Problem A has a feasible solution and that there exists a vector Z^o which satisfies Condition 3.6. Note that the specification of Z^o, assuming Condition 3.6, is trivial. (It is sufficient to define each of the k components of Z^o to be an arbitrary large positive number.) Perhaps the most significant weakness of this approach is the computation of the generalized inverse of an arbitrary matrix--this in itself is a significant computational task.

The computational aspects of this approach not withstanding, the development does provide insight with regards to the complexities of multiple criterion programming. In particular, this procedure provides a way to characterize the set of all efficient points by appealing to a subset $E_{\ell 2} \subseteq E$. This partitioning of efficient solutions based on the ℓ_2 matrix suggests other characterizations based on alternative metrics. Such is the motivation behind and primary thrust of the following chapter.

4. MINIMUM ℓ_p NORM PROBLEM AND CONVEX PROGRAMMING

In the previous chapter a linear version of the vector maximum problem was recast as a minimum ℓ_2 norm problem. This resulting problem was shown to be sufficient in the sense that a solution to the minimum norm problem defined a solution to the vector maximum problem. Since particular attention was focused on the Euclidean norm, it followed that the branch of convex nonlinear optimization, known as quadratic programming, was appealed to as the solution procedure. Moreover, an approach to the characterization of the set of all efficient solutions was presented based on the properties of the generalized inverse solution of a linear system. The insight developed with regard to the minimum ℓ_2 norm problem suggests a similar approach for the more general ℓ_p norm problem. Following in spirit the approach of Chapter 3, the primary emphasis of this chapter is focused on linear multiple objective programming problems. Within this context we will review and extend the concept of efficiency and construct a sufficient minimum norm problem. Since this more general problem is based on the minimum ℓ_p norm, it follows that a more general convex programming solution procedure be employed to identify the solutions of interest. To this end, it will be shown that a geometric programming problem can be constructed to serve our needs. Accordingly, a dual geometric programming formulation will be given with some rather extraordinary properties.

Before constructing the primal and dual geometric programs, let us restate the linear multiple objective programming problem as:

PROBLEM A

$$\text{"maximize"} \quad Cx$$

subject to

$$x \varepsilon S = \{x | x \varepsilon R^n, \ Ax \leq b, \ x \geq 0\}.$$

4.1 Another Classification of Efficiency

Recognize that a decision-maker might be overcome with the set of all efficient solutions to a linear multiple objective programming problem of the form given as Problem A. In Chapter 3 a procedure was presented whereby the decision-maker could characterize the set E with a subset $E_{\ell 2}$. Consider now, the more general concept of ℓ_p-efficient solutions and the resulting subset $E_{\ell p}$ to be used as a characterization of E. As with the definition of ℓ_2-efficiency, the definition of ℓ_p-efficiency depends critically on Condition 3.6 which we restate for convenience of reference as:

CONDITION 4.1 Assume there exists a vector $Z^o \varepsilon R^k$ such that $Z^o > Cx$ for all $x \varepsilon S = \{x | x \varepsilon R^n, \ Ax \leq b, \ x \geq 0\}$.

In view of the fact that the metric

$$\ell_p : \left[\sum_i |x_i|^p \right]^{1/p}$$

induces a norm, which we will denote as the ℓ_p norm $||.||p$, consider the following definition of ℓ_p-efficiency.

<u>DEFINITION 4.2</u> A point $x^o \varepsilon S$ is said to be ℓ_p-efficient if, and only if, there does not exist another point $x \varepsilon S$ such that

$$||Cx - Z^o||_p < ||Cx^o - Z^o||_p$$

where Z^o is a kxl vector which satisfies Condition 4.1.

For the sake of completeness it should be pointed out that the ℓ_p metric induces a norm provided that $p \geq 1$. For $p \varepsilon (0,1)$ the triangle inequality given in Definition 3.1 is reversed and, hence, the resulting function does not satisfy all of the required properties of a norm. In this case, we can view the ℓ_p norm as a "quasi-norm" (see [2]).

It can be shown (see [2]) that any norm defined by the ℓ_p metric, $p \varepsilon [1,\infty)$, is convex in the variable x--a very useful property that will be exploited in later sections. Throughout this chapter it is assumed that the norm, $||.||p$, is defined by the metric ℓp where $p \varepsilon [1,\infty)$.

As a consequence of Definition 4.2, we now introduce the following definition:

$$E_{\ell p} = \{x | x \varepsilon S \text{ and } x \text{ is } \ell_p\text{-efficient}\}.$$

Thus, a stronger version of Lemma 3.8 is now given as

<u>LEMMA 4.3</u> $E_{\ell p} \subseteq E$.

<u>Proof</u>. The proof of this lemma follows the proof of Lemma 3.8 where the Euclidean norm is replaced with the general ℓ_p norm.

Following in spirit the initial development of Chapter 3, consider the minimum norm problem:

<u>PROBLEM B</u> Find an x^o, if it exists, for which

$$||Cx^o - Z^o||_p = \underset{x \in S}{\text{minimum}} ||Cx - Z^o||_p$$

The relationship between this minimum norm problem and Problem A is given in the following result.

<u>LEMMA 4.4</u> Assume there exists an x^o which solves Problem B. Then

1. x^o is ℓ_p-efficient for Problem A, and
2. x^o is efficient for Problem A

<u>Proof</u>. The proof follows from Definition 4.2 and Lemma 4.3.

But Problem B may be expressed equivalently as:

PROBLEM C

$$\text{minimize} \quad ||\gamma||_p$$

subject to

$$||Cx - (Z^o - \gamma)||_p = 0$$

$$x \varepsilon S$$

$$\gamma \varepsilon R^k, \gamma \geqq 0.$$

Superficially, Problem C appears to be a more complex optimization problem. However, it will be shown that the resulting problem has a set of constraints that are virtually linear. That is, an equivalent set of linear constraints can be constructed as a substitute for the current set. Of more immediate interest are the following relationships between Problems A and C.

LEMMA 4.5 Problem A is feasible if, and only if, Problem C is feasible.

Proof. This proof follows the proof of Lemma 3.11 where the Euclidean norm is replaced with the more general ℓ_p norm.

LEMMA 4.6 If (x^o, γ^o) solves Problem C, then

1. x^o is ℓ_p-efficient for Problem A, and
2. x^o is efficient for Problem A.

Proof. The proof of this result follows as a consequence of Lemma 4.4 and the relationship between Problems B and C.

4.2 An Equivalent Geometric Program

The development of a solution procedure for ℓ_p-efficient solutions dictates a more generalized solution procedure than that presented in Section 3.2. In particular, since ℓ_2-efficiency is based on a minimum ℓ_2 distance, it followed that quadratic programming provided the computational support. However, we are now faced with a more general convex programming problem where geometric programming can be employed. To facilitate the construction of this geometric program we first note that

$$||\gamma||_p = \left[\sum_{i=1}^{k} |\gamma_i|^p \right]^{1/p} = \left[\sum_{i=1}^{k} \gamma_i^p \right]^{1/p}$$

since $\gamma \geq 0$. Moreover, it follows that

$$\text{minimize} \left[\sum_{i=1}^{k} \gamma_i^p \right]^{1/p}$$

is equivalent to

$$\text{minimize} \sum_{i=1}^{k} \gamma_i^p$$

since $p \geq 1$. With regard to the constraint set of Problem C we have

$$||Cx - (Z^o - \gamma)||_p = 0$$

if, and only if,

$$\left[\sum_{i=1}^{k} \left| \sum_{j=1}^{n} c_{ij}x_j - (Z_i^o - \gamma_i) \right|^p \right]^{1/p} = 0$$

or

$$\sum_{j=1}^{n} c_{ij}x_j + \gamma_i - Z_i^o = 0, \qquad i=1,\ldots,k.$$

Observe that this system may be expressed equivalently as

$$\sum_{j=1}^{n} c_{ij}x_j + \gamma_i - Z_i^o \leq 0, \qquad i=1,\ldots,k$$

$$-\sum_{j=1}^{n} c_{ij}x_j - \gamma_i + Z_i^o \leq 0, \qquad i=1,\ldots,k$$

In view of the above transformations we are now prepared to present an equivalent convex programming problem given as:

PROBLEM D

$$\text{minimize} \quad \sum_{i=1}^{k} \gamma_i^p$$

subject to

$$\sum_{j=1}^{n} c_{ij}x_j + \gamma_i - Z_i^o \leq 0 \qquad i=1,\ldots,k$$

$$-\sum_{j=1}^{n} c_{ij}x_j - \gamma_i + Z_i^o \leq 0 \qquad i=1,\ldots,k$$

$$\sum_{j=1}^{n} a_{ij}x_j - b_i \leq 0 \qquad i=1,\ldots,m$$

$$-x_j \leq 0 \qquad j=1,\ldots,n$$

$$-\gamma_i \leq 0 \qquad i=1,\ldots,k$$

To construct the geometric programming problem of interest we utilize the following one-to-one transformations (see Appendix C):

TRANSFORMATION L

$$x_j = \ln(w_j) \qquad j=1,\ldots,n$$

$$\gamma_i = \ln(w_{n+i}) \qquad i=1,\ldots,k$$

$$b_i = \ln(B_i) \qquad i=1,\ldots,m$$

$$Z_i^o = \ln(U_i) \qquad i=1,\ldots,k$$

The resulting convex programming problem will be of the form:

PROBLEM E (Primal Geometric Program)

minimize $G(w;p)$

subject to

$$U_i^{-1} \prod_{j=1}^{n} \left[w_j^{c_{ij}} \right] w_{n+i} \leq 1 \qquad i=1,\ldots,k$$

$$U_i \prod_{j=1}^{n} \left[w_j^{-c_{ij}} \right] w_{n+i}^{-1} \leq 1 \qquad i=1,\ldots,k$$

$$B_i^{-1} \prod_{j=1}^{n} w_j^{a_{ij}} \leq 1 \qquad i=1,\ldots,m$$

$$w_j^{-1} \lesssim 1 \qquad\qquad j=1,\ldots,n+k$$

with the implicit restriction that $w_j > 0$ for all j.

Here we make note of the fact that U_i, $B_i > 0$ for all i as a consequence of the logarithmic transformations employed. Moreover, if the objective function G is a posynomial then Problem E is amenable to the methods of geometric programming. For future reference we will let Ω define the set of all feasible solutions to Problem E. The task of interest is the construction of a posynomial G, which is a function of the vector w and the parameter p, such that if w^o solves Problem E then we can utilize our logarithmic transformations to define a solution x^o which solves Problem D.

In pursuit of this objective function $G(w;p)$ we first note that, ideally, we seek a posynomial G such that

$$\text{minimize} \sum_{i=1}^{k} \gamma_i^{\,p} \Leftrightarrow \text{minimize } G(w;p)$$

under Transformation L. Research to date has not yielded such a function. However, it is sufficient to identify a function G such that

$$\text{minimize } G(w;p) \Rightarrow \text{minimize} \sum_{i=1}^{k} \gamma_i^{\,p}$$

under Transformation L since this is sufficient to solve Problem D.

If we could define

$$G = \text{EXP}\left[\sum_{i=1}^{k} \gamma_i^{\,p}\right]$$

then we are done. However, such a function is not a prototype posynomial. Consider, now, the expression

$$\text{EXP}\left[\sum_{i=1}^{k} \gamma_i^{\,p}\right] = \text{EXP}\left[\sum_{i=1}^{k} (\ell n(w_{n+i}))^p\right]$$

where

$$\gamma_i = \ell n(w_{n+i}) \geq 0$$

where $w_{n+i} \geq 1$. Clearly,

$$\text{minimize} \quad \text{EXP}\left[\sum_{i=1}^{k} (\ell n(w_{n+i}))^p\right]$$

may be expressed equivalently as

$$\text{minimize} \quad \prod_{i=1}^{k} \text{EXP}\left[(\ell n(w_{n+i}))^p\right]$$

or

$$\text{maximize} \quad \prod_{i=1}^{k} \text{EXP}\left[-(\ell n(w_{n+i}))^p\right] .$$

But this is equivalent to

$$\text{maximize} \quad \ell n\{ \prod_{i=1}^{k} \text{EXP}\left[-(\ell n(w_{n+i}))^p\right] \}$$

or,

$$\text{maximize} \sum_{i=1}^{k} \ell n\{\text{EXP}\left[-(\ell n(w_{n+i}))^{p}\right]\}$$

which simplifies to

$$\text{maximize} \sum_{i=1}^{k} -(\ell n(w_{n+i}))^{p}$$

or, equivalently,

$$\text{minimize} \sum_{i=1}^{k} (\ell n(w_{n+i}))^{p}. \qquad (4.1)$$

Although this intermediate expression will, in theory, yield the desired result, the presence of the logarithmic function complicates the solution procedure. Further analysis of the functional suggests that we consider the relationship

$$w_{n+i} \geq \ell n(w_{n+i}), \qquad i=1,\ldots,k$$

which is true for any $w_{n+i} > 0$. Consequently, for $p \geq 1$ and $w_{n+i} \geq 1$,

$$w_{n+i}^{p} \geq (\ell n(w_{n+i}))^{p}, \qquad i=1,\ldots,k \qquad (4.2)$$

It then follows that an upper bound on the functional of interest in (4.1) is then

$$f_1(w;p) = \sum_{i=1}^{k} (\ell n(w_{n+i}))^{p} \leq \sum_{i=1}^{k} w_{n+i}^{p} = f_2(w;p). \qquad (4.3)$$

This intermediate result is of particular significance. Recognize first that if we define

$$G(w;p) = f_1(w;p)$$

then Problem E is completely equivalent to Problem D under Transformation L. If Ω defines the set of all feasible solutions to Problem E, then Problem E may be expressed as

$$\text{minimize} \quad f_1(w;p) \tag{4.4}$$

subject to

$$w \epsilon \Omega.$$

It is a well-known fact (see [41]) that w^o solves (4.4) if w^o solves the problem

$$\text{minimize} \quad f_2(w;p)$$

subject to

$$w \epsilon \Omega$$

and

$$f_1(w;p) \leq f_2(w;p) \tag{4.5}$$

But, in view of (4.2) and (4.3), expression (4.5) is <u>always</u> satisfied for <u>any</u> $w \epsilon \Omega$. Thus, we conclude at this point that it is sufficient to solve the following program.

PROBLEM F

$$\text{minimize} \quad \sum_{i=1}^{k} w_{n+i}^{p}$$

subject to

$$w \epsilon \Omega .$$

Observe now that the resulting program, Problem F, is amenable to the methods of geometric programming where the objective function is a posynomial and the constraint functions are single-term posynomials, or monomials. Moreover, as a consequence of our development, in particular expression (4.3), we can summarize our results thus far by stating the following.

LEMMA 4.7 If w^o solves Problem F, then there exists an x^o defined by Transformation L which solves Problem D.

Proof. If we define

$$G(w;p) = \sum_{i=1}^{k} (\ell n(w_{n+i}))^{p}$$

then, clearly, Problem E is completely equivalent to Problem D under Transformation L. Thus, the sufficiency of Problem F follows in view of the fact that expression (4.3) is valid for any $w\epsilon\Omega$.

Note that Problem F can be used to define a dual geometric programming problem (see Appendix B). This resulting dual problem is also amenable to the methods of convex (nonlinear) programming since the dual problem is one of maximizing a concave function over a set of linear (convex) constraints. However, in view of the relationship between linear programming and geometric programming with monomial functions (see Appendix C), it behooves us to consider the feasibility of yet another surrogate objective function which is monomial.

Consider specifying $G(w;p)$ to be a monomial function of the vector w and the parameter p such that

$$\sum_{i=1}^{k} w_{n+i}^{p} \leq G(w;p).$$

Assuming such a monomial function exists, this suggests the following extension of Problem F.

PROBLEM G

$$\text{minimize} \quad G(w;p)$$

subject to

$$w \varepsilon \Omega$$

and

$$\sum_{i=1}^{k} w_{n+i}^{p} \leq G(w;p) \tag{4.6}$$

To demonstrate that such a monomial function always exists we have:

LEMMA 4.8 Let $u_1, u_2, \ldots, u_n$ be real numbers that satisfy $u_i \geq 1$ for $i=1,\ldots,n$. Then

$$\sum_{i=1}^{n} u_i^p \leq n \prod_{i=1}^{n} u_i^p$$

for any $p \in [1,\infty)$.

Proof. Observe that

$$n \prod_{i=1}^{n} u_i^p = \sum_{i=1}^{n} \prod_{i=1}^{n} u_i^p. \tag{4.7}$$

Now, to establish the validity of this lemma it suffices to show that

$$u_i^p \leq \prod_{i=1}^{n} u_i^p \quad \text{for} \qquad j=1,\ldots,n.$$

But this follows immediately since $u_i \geq 1$ for all i and $p \in [1,\infty)$.

As a consequence of Lemma 4.8, if we now define

$$G(w;p) = k \prod_{i=1}^{k} w_{n+i}^p$$

then Problem G takes the form

$$\underset{w \in \Omega}{\text{minimize}} \; k \prod_{i=1}^{k} w_{n+i}^p .$$

Note that explicit inclusion of expression (4.6) is unnecessary in view of the fact that $w \in \Omega$ implies that $w \geq 1$. Hence, expression (4.6) will always be satisfied, in view of Lemma 4.8, which suggests that a solution to Problem D may be obtained by solving the program given as:

PROBLEM H

$$\text{minimize} \quad G(w;p) = k \prod_{i=1}^{k} w_{n+i}^{p}$$

subject to

$$U_i^{-1} \prod_{j=1}^{n} \left[w_j^{c_{ij}} \right] w_{n+i} \leq 1 \qquad i=1,\ldots,k$$

$$U_i \prod_{j=1}^{n} \left[w_j^{-c_{ij}} \right] w_{n+i}^{-1} \leq 1 \qquad i=1,\ldots,k$$

$$B_i^{-1} \prod_{j=1}^{n} w_j^{a_{ij}} \leq 1 \qquad i=1,\ldots,m$$

$$w_j^{-1} \leq 1 \qquad j=1,\ldots,n+k.$$

Before investigating the corresponding dual geometric programming we will make some definitions to obtain notational simplicity. Since each function has a single term (monomial functions), it follows that each function will have one associated coefficient. Thus, we incorporate the substitutions:

$$D_o = k$$

$$D_i = U_i^{-1} \qquad i=1,\ldots,k$$

$$D_{k+i} = U_i \qquad i=1,\ldots,k$$

$$D_{2k+i} = B_i^{-1} \qquad i=1,\ldots,m$$

$$D_{2k+m+i} = 1 \qquad i=1,\ldots,n+k$$

Likewise, for the exponents of the variables, w_j, we define:

$$E = \left[\begin{array}{c|c} 0_{1,n} & (p,\ldots,p)_{1,k} \\ \hline C_{k,n} & I_{k,k} \\ \hline -C_{k,n} & -I_{k,k} \\ \hline A_{m,n} & 0_{m,k} \\ \hline \multicolumn{2}{c}{-I_{(n+k),(n+k)}} \end{array}\right]$$

so that $E_{M,N}$ is $(3k + m + n + 1) \times (n + k)$. Here the index on the element e_{ij} when $i = 0$ denotes the jth element in the first row of E. Accordingly, Problem H may be simplified notationally and expressed as:

<u>PROBLEM H'</u>

$$\text{minimize } D_o \prod_{j=1}^{N} w_j^{e_{oj}}$$

subject to

$$D_i \prod_{j=1}^{N} w_j^{e_{ij}} \le 1 \qquad i=1,\ldots,M$$

$$w_j > 0 \qquad j=1,\ldots,N$$

Clearly, Problem H' is a geometric programming problem with monomial functions throughout. Summarizing our development thus far, we have the following intermediate result.

LEMMA 4.9 If w^o solves Problem H, then there exists an x^o defined by Transformation L which solves Problem D.

Proof. Lemma 4.7 establishes the fact that if w^o solves Problem F then Transformation L can be used to define a solution x^o which solves Problem D. Thus, the validity of this result is established by showing that if w^o solves Problem H then w^o solves Problem F. But the sufficiency of Problem H for Problem F follows as a consequence of Lemma 4.8 which implies that

$$\sum_{i=1}^{k} w_{n+i}^{p} \leq k \prod_{i=1}^{k} w_{n+i}^{p}$$

when $w \in \Omega$.

4.3 A Dual Problem

Consider, for a moment, the ramifications of Lemma 4.9. The geometric program, Problem H, is not completely equivalent to the minimum ℓ_p norm program of interest-Problem D. However, Problem H is sufficient for Problem D in the sense that a solution to Problem H can be used, in view of Transformation L, to define a solution to Problem D. From a computational perspective, this result appears to be of little value since the original problem of interest is a linearly constrained minimization of a convex function. Clearly, the highly nonlinear nature of Problem H indicates that a sophisticated procedure be employed to obtain the solution. However, as a consequence of duality in geometric programming, we are now in a position to construct a dual problem which exhibits significant computational advantages. Moreover, as a consequence of the first and second duality theorems of geometric programming (Theorems B.2 and B.3, respectively, in Appendix B), the resulting dual problem is completely equivalent to its corresponding primal, Problem H. The dual geometric programming problem corresponding to Problem H' is given as:

PROBLEM I Find a vector, δ^o, such that

$$v(\delta^o) = \text{minimum } v(\delta)$$

where

$$v(\delta) = \prod_{i=0}^{M} \left[\left(\frac{D_i}{\delta_i} \right)^{\delta_i} \right] \prod_{i=0}^{M} \lambda_i(\delta)^{\lambda_i(\delta)}$$

subject to

$$\lambda_o(\delta) = 1 \qquad \text{(normality)}$$

$$\sum_{i=0}^{M} e_{ij}\,\delta_i = 0 \qquad j=1,2,\ldots,N \qquad \text{(orthogonality)}$$

$$\delta_i \geqq 0, \qquad i=0,1,\ldots,M \qquad \text{(positivity)}$$

Here $\lambda_i(\delta) = \delta_i$ for $i=0,1,\ldots M$ as presented in Appendix B.

Observe that the resulting dual problem is linearly constrained but that the objective function, $v(\delta)$, is quite complex and highly non-linear. With this problem as with the corresponding primal problem, Problem H, the computational advantages are, at best, questionable. Fortunately, the objective function can be simplified significantly to

$$v(\delta) = \prod_{i=0}^{M} \left[\left(\frac{D_i}{\delta_i} \right)^{\delta_i} \right] \prod_{i=1}^{M} \lambda_i(\delta)^{\lambda_i(\delta)}$$

$$= D_o \prod_{i=1}^{M} \left[\left(\frac{D_i}{\delta_i} \right)^{\delta_i} (\delta_i)^{\delta_i} \right]$$

$$= D_o \prod_{i=1}^{M} D_i^{\delta_i} .$$

Also, the monotonicity of the logarithmic functions (see [16]) guarantees that

maximize $v(\delta)$

is completely equivalent to

maximize $V(\delta)$

where

$$V(\delta) = \ell n(v(\delta)) = \ell n \left[D_o \prod_{i=1}^{M} D_i^{\delta_i} \right] .$$

Thus, the dual objective function can be expressed equivalently as

$$V(\delta) = \ell n(D_o) + \sum_{i=1}^{M} \ell n(D_i)\delta_i.$$

Here, of course, $D_i > 0$ for all i as a consequence of Problem E. That is, the D_i are coefficients for the monomial functions in Problem E which are guaranteed to be strictly positive. Reverting back to these original coefficients and applying Transformation L, it can easily be shown that the dual objective function assumes the form

$$V(\delta) = \ell n(k) + \sum_{i=1}^{k} - Z_i^o\delta_i + \sum_{i=1}^{k} Z_i^o\delta_{k+i} + \sum_{i=1}^{m} - b_i\delta_{2k+i}$$

$$+ \sum_{i=2k+m+1}^{3k+m+n} \ell n(1)\delta_i$$

Omitting constant terms and simplifying $V(\delta)$ we have that the dual problem, expressed in terms of the original coefficients, is given as

PROBLEM J

$$\text{maximize} \quad \sum_{i=1}^{k} - Z_i^o\delta_i + \sum_{i=1}^{k} Z_i^o\delta_{k+i} + \sum_{i=1}^{m} -b_i\delta_{2k+i}$$

subject to

$$\left[\begin{array}{c:c:c} C^T_{n,k} & -C^T_{n,k} & A^T_{n,m} \\ \hdashline I_{k,k} & -I_{k,k} & 0_{k,m} \end{array} \quad -I\right] \begin{bmatrix} \delta_1 \\ \delta_2 \\ \vdots \\ \delta_{3k+m+n} \end{bmatrix} = \begin{bmatrix} \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}_n \\ \begin{pmatrix} -p \\ \vdots \\ -p \end{pmatrix}_k \end{bmatrix}$$

and

$$\delta_i \geq 0, \qquad i=1,2,\ldots,3k+m+n.$$

As a consequence of the first duality theorem of geometric programming, (Appendix B, Theorem B.2), we can recover the optimal solution to the primal problem, given that δ^o solves the dual problem, by solving the system:

$$D_o \prod_{j=1}^{N} (w_j^o)^{e_{oj}} = v(\delta^o)$$

$$D_i \prod_{j=1}^{N} (w_j^o)^{e_{ij}} = 1 \qquad i=1,\ldots,M \text{ for which } \delta_i^o > 0.$$

Or, equivalently, we can determine $\ln(w_j^o)$ $j=1,\ldots,N$ by solving the linear system:

$$\ln(D_o) + \sum_{j=1}^{N} e_{oj} \ln(w_j^o) = \ln(v(\delta^o))$$

$$\ln(D_i) + \sum_{j=1}^{N} e_{ij} \ln(w_j^o) = 0, \quad i=1,\ldots,M \text{ for which } \delta_i^o > 0.$$

The solution to the linear programming problem, Problem J, is of particular significance as indicated by the following result.

LEMMA 4.10 If there exists a $\delta^o \epsilon R^{3k+m+n}$ which is an optimal solution to Problem J, then

1. there exists some w^o which solves Problem H,

2. $x_j^o = \ell n(w_j^o)$, j-1,...,n+k solves Problem D, and

3. $x^o = (x_1^o, x_2^o, \ldots, x_n^o)'$ is ℓ_p-efficient for Problem A.

Proof. (1) follows as a consequence of the theory of duality in geometric programming (see Lemma B.1, Theorems B.2 and B.3 of Appendix B). Lemma 4.9 and Transformation L, under (1), then imply (2). Finally, the equivalence of Problems C and D together with Lemma 4.6 imply (3).

Before proceeding further with extensions of this approach, it is instructive to summarize the procedure developed thus far. Reconsider the linear multiple objective programming problem of the form given in Problem A. Under Condition 4.1, we can construct a minimum ℓ_p norm problem, Problem C, the solution of which is, by definition, ℓ_p-efficient for Problem A. Utilizing a sequence of logarithmic transformations and upper bounding inequalities, we have constructed a geometric programming problem which is "computationally sufficient" in the sense that it identifies solutions to the minimum ℓ_p norm problem, Problem C. An application of duality and additional transformations result in an equivalent dual problem which can be solved directly by linear programming techniques. The flow diagram in Figure 1 describes this equivalence chain.

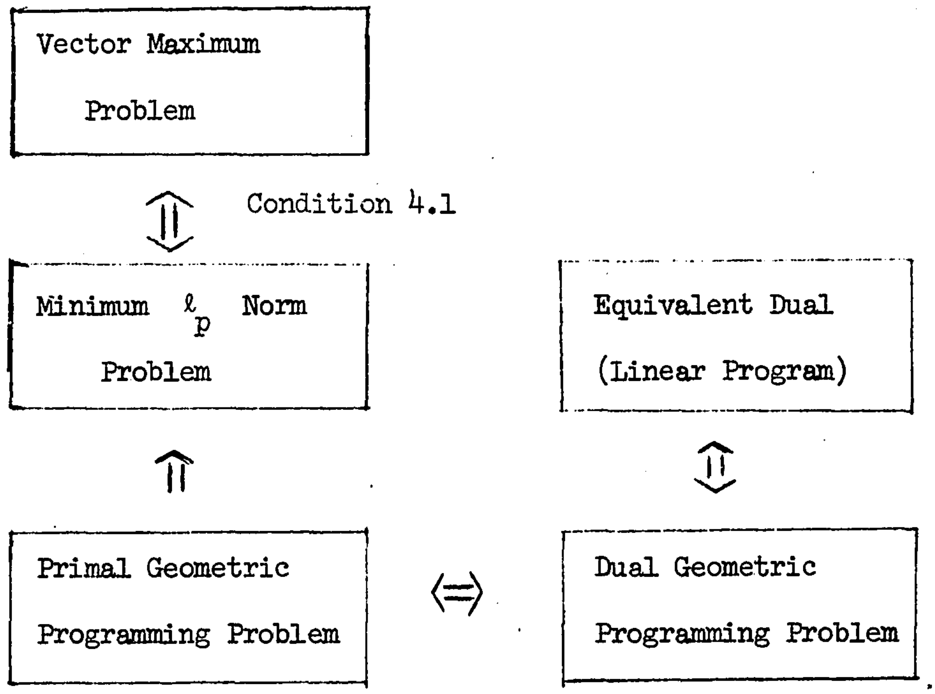

Figure 1. Equivalence Structure With Condition 4.1

4.4 Extensions For A More General Problem

In this section we present the machinery needed to extend the results of the previous sections. In particular, we will focus attention on a more general problem and, effectively, sharpen the results presented thus far. Recall that the minimum ℓ_p norm program, Problem D, is sufficient for Problem A only under the rather strong assumption given as Condition 4.1. Motivation for the development of a procedure which permits the relaxation of Condition 4.1 is based on the potential application of these results to such problem areas as linear regression and goal programming. As presented in Chapter 2, goal programming, to date, has been concerned with maximizing a measure of goal achievement (achievement is assumed to be synonymous with utility) where the particular measure is based on the familiar ℓ_1 metric. If indeed one could relax Condition 4.1, then this procedure could be applied to what may be termed generalized goal programming or, perhaps, convex goal programming as recently introduced by Charnes and Cooper [7] and studied by Charnes, Cooper, Klingman, and Niehaus [8]. Let us revisit Condition 4.1 and consider the implications of this assumption with regard to a more general problem.

Recall that Problem A has a constraint set of the form

$$S = \{x | x \varepsilon R^n, Ax \leqq b, x \geqq 0\}.$$

Condition 4.1 states that there exists a $k x l$ vector Z^o such that $Z^o > Cx$ for all $x \varepsilon S$. In the context of a goal programming problem (see Section 2.2), this condition is clearly unwanted. For example, a generalized goal programming problem could be formulated as

$$\underset{x \varepsilon S}{\text{minimize}} \quad ||Cx - Z^o||_p$$

where $Z^o \varepsilon R^k$ is an arbitrary vector not subject to the assumption that $Z^o > Cx$ for all $x \varepsilon S$ since the goal formulation here is one of simultaneous attainment as opposed to simultaneous maximization. Such is also the formulation of constrained regression problems which we will consider further in the next chapter. Without loss of generality the more general problem may be stated as

PROBLEM K

$$\text{minimize} \quad ||Cx - Z^o||_p$$

subject to

$$x \varepsilon S = \{x | Ax \lesseqgtr b\}$$

where $p \varepsilon [1, \infty)$ and $Z^o \varepsilon R^k$ is arbitrary.

Observe that we have redefined S so that explicit restriction of the vector x to the nonnegative orthant is relaxed. Certainly, some or all of the components of x could be constrained to be nonnegative within the new formulation of S if desired. The significance difference in this formulation is that Z^o is arbitrary so that Condition 4.1 is not appropriate. Before examining the ramifications of Condition 4.1 and the more general problem we follow, in spirit, the transformations presented earlier in this chapter and expresses Problem K equivalently as

PROBLEM M

$$\text{minimize} \quad ||\gamma||_p$$

subject to

$$||Cx - (Z^o - \gamma)||_p = 0$$
$$Ax \leq b$$
$$x\epsilon R^n, \quad \gamma\epsilon R^k.$$

Since $Z^o \epsilon R^k$ is now arbitrary, it follows that the deviation vector $\gamma\epsilon R^k$ must have components which are unrestricted in sign. This is precisely the motivation behind Condition 4.1 in that this assumption insures that

$$Z^o - Cx = \gamma \geq 0$$

which is essential in the proof of Lemma 4.5. For the more general problem a substitution of variables can be used to circumvent this dilemma so that the results presented thus far are valid without enforcing Condition 4.1. Let us re-state Problem M as

$$\text{minimize} \; ||\gamma||_p$$
$$(x,\gamma)\epsilon\Omega,$$

where Ω denotes the set of all feasible solutions to Problem M, and note that this program may be expressed equivalently as

$$\text{minimize} \quad ||y||_p$$

subject to

$$||\gamma||_p \leq ||y||_p \tag{4.8}$$
$$(x,\gamma)\varepsilon\Omega$$
$$y\varepsilon R^k, \; y \geq 0.$$

Note further that restriction (4.8) is guaranteed to be satisfied when the restriction

$$-y_i \leq \gamma_i \leq y_i \qquad i=1,\ldots,k$$

is enforced. Applying these substitutions we have that

$$\text{minimize} \quad ||y||_p$$

subject to

$$-y_i \leq \gamma_i \leq y_i \qquad i=1,\ldots,k$$
$$(x,\gamma)\varepsilon\Omega$$
$$y\varepsilon R^k, \; y \geq 0$$

is sufficient for

$$\text{minimize} \quad ||\gamma||_p.$$

subject to

$$(x, \gamma)\,\varepsilon\Omega$$

Here, of course, $p \epsilon [1,\infty)$. Thus, a counterpart to Problem D which is sufficient for the more general problem, Problem M, is given as

PROBLEM N

$$\text{minimize} \quad ||y||_p$$

subject to

$$\sum_{j=1}^{n} c_{ij}x_j + \gamma_i - Z_i^o \leq 0 \qquad i=1,\ldots,k$$

$$\sum_{j=1}^{n} - c_{ij}x_j - \gamma_i + Z_i^o \leq 0 \qquad i=1,\ldots,k$$

$$\sum_{j=1}^{n} a_{ij}x_j - b_i \leq 0 \qquad i=1,\ldots,m$$

$$\gamma_i - y_i \leq 0 \qquad i=1,\ldots,k$$

$$-\gamma_i - y_i \leq 0 \qquad i=1,\ldots,k$$

$$-y_i \leq 0 \qquad i=1,\ldots,k$$

$$\gamma \epsilon R^k, \; x \epsilon R^n$$

This resulting formulation of Problem N is identical, in construction, to the linearly constrained ℓ_p norm program given as Problem D since nonnegativity of the vector y ensures that

$$\text{minimize} \quad ||y||_p$$

is equivalent to

$$\text{minimize} \sum_{i=1}^{k} y_i^p \, .$$

It then follows that a sufficient geometric program with monomial functions can now be formulated if we redefine our logarithmic transformations to be

TRANSFORMATION L'

$$\begin{aligned} x_i &= \ln(w_j) & i&=1,\ldots,n \\ \gamma_i &= \ln(w_{n+i}) & i&=1,\ldots,k \\ y_i &= \ln(w_{n+k+i}) & i&=1,\ldots,k \\ b_i &= \ln(B_i) & i&=1,\ldots,m \\ Z_i^o &= \ln(U_i) & i&=1,\ldots,k \end{aligned}$$

Utilizing our new transformation, Transformation L', the geometric program can be expressed as

PROBLEM O

$$\text{minimize } k \prod_{j=1}^{k} w_{n+k+j}^p$$

subject to

$$U_i^{-1} \prod_{j=1}^{n} \left[w_j^{c_{ij}} \right] w_{n+i} \leq 1 \qquad i=1,\ldots,k$$

$$U_i \prod_{j=1}^{n} \left[w_j^{-c_{ij}} \right] w_{n+i}^{-1} \leq 1 \qquad i=1,\ldots,k$$

$$B_i^{-1} \prod_{j=1}^{n} w_j^{a_{ij}} \leq 1 \qquad i=1,\ldots,m$$

$$w_{n+i}\, w_{n+k+i}^{-1} \leq 1 \qquad i=1,\ldots,k$$

$$w_{n+i}^{-1}\, w_{n+k+i}^{-1} \leq 1 \qquad i=1,\ldots,k$$

$$w_{n+k+i}^{-1} \leq 1 \qquad i=1,\ldots,k$$

with the implicit restriction that $w_i > 0$ for $i=1,\ldots,n+2k$.

Observe that Problem O is identical in construction to the geometric program given as Problem H. Therefore, it follows that the substitutions for coefficients:

$$D_o = k$$

$$D_i = U_i^{-1} \qquad i=1,\ldots,k$$

$$D_{k+i} = U_i \qquad i=1,\ldots,k$$

$$D_{2k+i} = B_i^{-1} \qquad i=1,\ldots,m$$

$$D_{2k+m+i} = 1 \qquad i=1,\ldots,3k$$

and the substitution for exponents of the variables w_j:

$$E = \begin{bmatrix} O_{1,n} & O_{1,k} & (p,\ldots,p)_{1,k} \\ C_{k,n} & I_{k,k} & O_{k,k} \\ -C_{k,n} & -I_{k,k} & O_{k,k} \\ A_{m,n} & O_{m,k} & O_{m,k} \\ O_{k,n} & I_{k,k} & -I_{k,k} \\ O_{k,k} & -I_{k,k} & -I_{k,k} \\ O_{k,k} & O_{k,k} & -I_{k,k} \end{bmatrix}_{M,N}$$

permit Problem 0 to be expressed more conveniently as

$$\text{minimize} \quad D_o \prod_{j=1}^{N} w_j^{e_{ij}}$$

subject to

$$D_i \prod_{j=1}^{N} w_j^{e_{ij}} \leq 1 \qquad i=1,\ldots,M$$

$$w_j > 0 \qquad j=1,\ldots,N.$$

Accordingly, the resulting dual geometric program is of the form:

PROBLEM P

$$\text{maximize} \quad \prod_{i=0}^{M} \left[\left(\frac{D_i}{\delta_i} \right)^{\delta_i} \right] \prod_{i=1}^{M} \lambda_i(\delta)^{\lambda_i(\delta)}$$

subject to

$$\lambda_o\,(\delta) = 1 \qquad \text{(normality)}$$

$$\sum_{i=0}^{M} e_{ij}\delta_i = 0, \qquad j=1,\ldots,N \qquad \text{(orthogonality)}$$

$$\delta_i \geqq 0, \qquad i=1,\ldots,M \qquad \text{(positivity)}$$

when $\lambda_i\,(\delta) = \delta_i$ for $i=0,1,\ldots,M$ as presented in Appendix B.

As with Problem J, this dual problem can be simplified considerably and expressed equivalently as:

PROBLEM Q

$$\text{maximize} \quad \sum_{i=1}^{k} - Z_i^o\delta_i + \sum_{i=1}^{k} Z_i^o\delta_{k+i} + \sum_{i=1}^{m} - b_i\delta_{2k+i}$$

subject to

$$\begin{bmatrix} C^T_{n,k} & -C^T_{n,k} & A^T_{n,m} & 0_{n,k} & 0_{n,k} & 0_{n,k} \\ I_{k,k} & -I_{k,k} & 0_{k,m} & I_{k,k} & -I_{k,k} & 0_{k,k} \\ 0_{k,k} & 0_{k,k} & 0_{k,m} & -I_{k,k} & -I_{k,k} & -I_{k,k} \end{bmatrix} \begin{bmatrix} \delta_1 \\ \delta_2 \\ \cdot \\ \cdot \\ \cdot \\ \delta_{5k+m} \end{bmatrix} = \begin{bmatrix} \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}_n \\ \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}_k \\ \begin{pmatrix} -p \\ \vdots \\ -p \end{pmatrix}_k \end{bmatrix}$$

and

$$\delta_i \geq 0, \qquad i=1,2,\ldots,5k+m.$$

The solution procedure, when extended to deal with a more general problem of the form

$$\text{minimize} \quad ||Cx - Z^o||_p$$

subject to

$$x \varepsilon S = \{x | x \varepsilon R^n, \ Ax \leq b,\}$$

where $p \varepsilon [1,\infty)$ and $Z \varepsilon R^k$ is arbitrary, is potentially a significant computational result. In its full generality it permits such a convex nonlinear optimization problem to be solved by linear programming techniques. Indeed the key results of this section on extensions are that $Z^o \varepsilon R^k$ can be arbitrary and that the decision vector x need not be constrained into the nonnegative orthant. Moreover, the relaxation of Condition 4.1 permits the exclusion or inclusion of the set S to be optional. These extensions, therefore, expand the scope of the procedure to include the generalized linear regression problem (referred to as the ℓ_p approximation problem in Chapter 2) with or without linear side conditions as well as the generalized goal programming problem based on the ℓ_p metric. This equivalence chain is described in Figure 2.

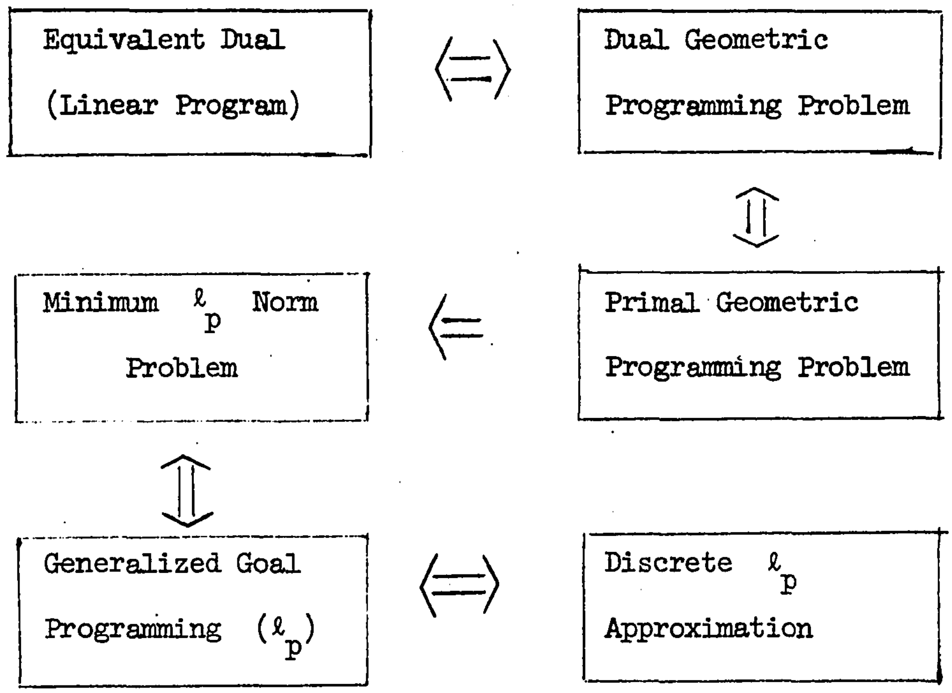

Figure 2. Equivalence Structure for $p \in [1,\infty)$

5. ASPECTS OF DUALITY

In the previous chapter, a sufficient programming problem was constructed to solve a linearly constrained minimum ℓ_p norm problem. Furthermore, duality was employed to provide an equivalent problem, to the sufficient program, with some rather significant computational advantages. Note that the theory of duality was utilized for the resulting (sufficient) geometric programming problem and not the minimum norm problem itself. The results of Chapter 4 not withstanding, this suggests the plausibility of applying duality results directly to the minimum norm problem in pursuit of a more computationally attractive model.

The intermediate results presented in this chapter do not contribute significantly to the computational aspects of the problem. However, in the spirit of Chapter 3, this development is intended to provide insight with regard to alternative dual problems associated with linearly constrained minimum ℓ_p norm problems. The motivation for the inclusion of this work is derived from the fact that a constrained minimum norm optimization problem is fundamental to such application areas as linear regression and generalized goal programming based on the ℓ_p metric.

In considering dual problems we will note that several approaches to duality are prominent in the literature of mathematical programming. In particular, the derivation of the dual geometric program, presented in Appendix B, is based on a problem defined over a vector space which is the orthogonal complement of the primal space. Hence, in this case, duality is based on the orthogonality of two distinct vector spaces. Perhaps the most common approach to the construction and study of dual problems is based on the Lagrangian function (see Kuhn and Tucker [33], Mangasarian [37] or Sposito [41]) which we shall address as Lagrangian duality. (It is interesting to note that the Lagrangian approach and the orthogonality approach to duality in geometric programming result in identical dual problems as evidenced by Theorem B.2 of Appendix B.) We now present some relevant results on duality for the linearly constrained minimum norm problem.

5.1 Orthogonal Duality

The results of the previous chapters are based on the norm induced by the ℓ_p metric where $p \geq 1$. Recall that a norm is any real-valued function which satisfies the properties of a norm. In pursuit of generality, it will be assumed that the norm $||.||$ denotes <u>any</u> functional satisfying axioms 1 through 4 of Definition 3.1. Thus we include, as a special case, the ℓ_p norm; but we are certainly not restricting ourselves to this function. With regard to the orthogonality of vector spaces, we first introduce the definition of alignment.

<u>DEFINITION 5.1</u> A vector $x \varepsilon X$ is said to be aligned with a vector $y \varepsilon Y$ if

$$\langle x,y\rangle = ||x||.||y||.$$

Observe that alignment is a relationship between two vectors in two distinct vector spaces. In this case, the vector spaces are the normal space X and its normal dual space Y.

Consider, now, the linearly constrained minimum norm problem given as:

PROBLEM A

$$\text{minimize} \quad ||x||$$

subject to

$$x \varepsilon S = \{x | x \varepsilon R^n, \ Ax = b\}$$

Following the development presented in Luenberger [36], let x^* be any vector satisfying the constraints of Problem A. Then we have that

$$d = \underset{x \varepsilon S}{\text{minimum}} \ ||x||$$

$$= \underset{y \varepsilon \bar{Y}}{\text{minimum}} \ ||x^* - y||.$$

Here Y denotes the space generated by the rows of the matrix A and, accordingly, Y denotes the orthogonal complement of Y (i.e., $\bar{Y} = \{y | \langle y,z\rangle = 0, \ z \varepsilon Y\}$). It then follows (by an application of Theorem 2, page 121 in [36]) that

$$d = \underset{y \in \bar{Y}}{\text{minimum}} \; ||x^* - y|| = \underset{\substack{x \in Y \\ ||x|| \leq 1}}{\text{supremum}} \langle x, x^* \rangle .$$

Any vector in Y is of the form

$$x = \sum_{i=1}^{m} w_i a_i$$

where a_i represents the ith row of the matrix A and $w \in R^m$. We represent this situation notationally as $A'w$. Thus, since Y is of finite dimension,

$$d = \underset{x \in S}{\text{minimum}} \; ||x|| = \underset{||A'w|| \leq 1}{\text{maximum}} \langle x, x^* \rangle$$

$$= \underset{||A'w|| \leq 1}{\text{maximum}} \langle A'w, x^* \rangle$$

$$= \underset{||A'w|| \leq 1}{\text{maximum}} \; b'w$$

where the last equality follows from the fact that x^* satisfies the constraints of Problem A (i.e., $Ax^* = b$). The results of this analysis are summarized in the following corollary to the development.

<u>COROLLARY 5.2 (Luenberger [36])</u> Let the linear system

$$S = \{x | x \in R^n, \; Ax = b\},$$

where A is mxn and b is mx1, be nonempty. Then

$$\underset{x\varepsilon S}{\text{minimum}}\ ||x|| = \underset{||A'w||\leq 1}{\text{maximum}}\ b'w.$$

Moreover, the optimal x^o is aligned with the optimal $A'w^o$ so that

$$\langle x^o, A'w^o\rangle = ||x^o||\cdot||A'w^o||.$$

Although Corollary 5.2 states an interesting theoretical result, the corresponding dual problem does not, apparently, afford any significant computational advantage relative to the original primal program, Problem A. However, the development of this dual problem, in particular the property of alignment of the optimal vectors in their corresponding dual spaces, suggests the following approach to duality via the Lagrangian function.

5.2 Lagrangian Duality

Analysis of dual problems derived from the Lagrangian function has proved beneficial in optimization theory. Moreover, it is sometimes the case that a dual problem constructed from the Lagrangian function has significant computational advantages as with linear and quadratic programming (see Mangasarian [37] or Sposito [41]). Although research to date has not yielded significant results for dual problems associated with the general minimum ℓ_p norm problem, it is instructive to study the relationship between the primal problem,

Problem A, and a corresponding dual problem given as:

PROBLEM B

$$\text{maximize} \quad b'w$$

subject to

$$||A'w|| \leq 1.$$

Consider the Lagrangian function associated with Problem A expressed as

$$\phi(x,w) = ||x|| + w'(b - Ax) \tag{5.1}$$

which is defined over $x \epsilon S$ and $w \epsilon R^m$. Likewise, for the dual problem, Problem B, we have

$$\psi(w,x) = b'w + g(x)\ (1 - ||A'w||) \tag{5.2}$$

Observe that (5.2) includes a function $g(x)$ which may be viewed as a Lagrangian multiplier in the same sense that the vector w' relates to expression (5.1) (i.e., w is a vector of Lagrangian multipliers or dual variables). It is interesting to note that the function $g(x)$ must provide a mapping $g:R^n \to R$ in view of the fact that Problem B has only one constraint. Research on the relationship between Problem B and expression (5.2) indicates that if we define the function $g(x)$ to be

$$g(x) = ||x||$$

then we can verify the relationship between (5.1) and (5.2). In support of Corollary 5.2, we assert the following result.

PROPOSITION 5.3 If x^o is an optimal solution to (primal) Problem A and w^o is an optimal solution to (dual) Problem B, then

$$\phi(x^o,w^o) = \psi(w^o,x^o).$$

Proof. Upon rearranging terms in (5.1), we have that

$$\begin{aligned} \phi(x^o,w^o) &= ||x^o|| + w^{o\prime}(b - Ax^o) \\ &= ||x^o|| + w^{o\prime}b - w^{o\prime}Ax^o \\ &= b'w^o + ||x^o|| - x^{o\prime}A'w^o. \end{aligned} \tag{5.3}$$

Likewise, substituting $g(x) = ||x||$ into (5.2) we have

$$\begin{aligned} \psi(w^o,x^o) &= b'w^o + ||x^o||(1 - ||A'w^o||) \\ &= b'w^o + ||x^o|| - ||x^o|| \cdot ||A'w^o||. \end{aligned} \tag{5.4}$$

But, by the alignment property of optimal x^o and optimal $A'w^o$ (Definition 5.1 and Corollary 5.2), it follows that

$$||x^o|| \cdot ||A'w^o|| = \langle x^o, A'w^o \rangle = x^{o\prime}A'w^o$$

which implies, in view of (5.3) and (5.4), that $\phi(x^o,w^o) = \psi(w^o,x^o)$.

The fact that Proposition 5.2 holds is not, in itself, a significant result. Actually, if it were not true that $\phi(x^o,w^o) = \psi(w^o,x^o)$ then one would have just cause for questioning the validity of the dual program, Problem B. The interesting result is that the Lagrangian multiplier utilized in the function $\psi(w,x)$ takes the form of norm $||x||$. In mathematical programming, the dual variables (Lagrangian multipliers) are often interpreted as "shadow costs" (see [37], [41]). Further analysis of the multiplier $||x||$ in the context of a shadow cost and interpretation of the physical significance of this functional would indeed be an interesting area for future investigation. Moreover, in the context of a generalized goal programming problem (based on the ℓ_p metric), analysis of the dual problem, Problem B, in view of expression (5.2) might identify those aspects of duality, indifference, and sensitivity analysis not covered in the duality results by Kornbluth [31].

6. CONCLUSIONS AND RECOMMENDATIONS FOR FURTHER RESEARCH

The contribution of this thesis may be characterized as providing a computationally attractive approach to the linearly constrained minimum ℓ_p norm problem. Inasmuch as the minimum norm problem is fundamental to such problem areas as linear multiple objective programming and linear regression (with or without constraints), it follows that the results of this study provide a marginal contribution to these areas as well.

Although we have used the adjective "attractive" to describe the procedure presented herein, a valid criticism can be made regarding the size of the resulting dual problem--albeit a linear model. In particular, we will now consider the dual problem and suggest a procedure to expedite the solution.

6.1 Reduction of the Working Basis

With regard to the computational aspects of the problem, let us revisit the dual problem, Problem Q, in Chapter 4. Observe, in particular, the (linear) constraints on the model. It is clear that the size of this linear program can become quite large as the number of decision variables, constraints, and goal functions increase. Although this approach affords the advantage of linear optimization, it is evident that a multiple criterion program of moderate size requires the solution of a dual problem (i.e., Problem Q) which is bordering on a "large-scale programming" problem. Analysis of the model suggests that we consider equivalent formulations of this problem to expedite the solution procedure. In particular, the linear system

$$\begin{bmatrix} C^T_{n,k} & -C^T_{n,k} & A^T_{n,m} & 0_{n,k} & 0_{n,k} & 0_{n,k} \\ I_k & -I_k & 0_{k,m} & I_k & -I_k & 0_{k,k} \\ 0_{k,k} & 0_{k,k} & 0_{k,m} & -I_k & -I_k & -I_k \end{bmatrix} \delta = \begin{bmatrix} \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}_n \\ \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}_k \\ -\begin{pmatrix} p \\ \vdots \\ p \end{pmatrix}_k \end{bmatrix}$$

$$\delta_1, \delta_2, \ldots, \delta_{5k+m} \geq 0$$

can be expressed equivalently as

$$\begin{bmatrix} C^T_{n,k} & A^T_{n,m} & 0_{n,k} & 0_{n,k} & 0_{n,k} \\ I_k & 0_{k,m} & I_k & -I_k & 0_{n,k} \\ 0_{k,k} & 0_{k,m} & I_k & I_k & I_k \end{bmatrix} y = \begin{bmatrix} \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}_n \\ \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}_k \\ \begin{pmatrix} p \\ \vdots \\ p \end{pmatrix}_k \end{bmatrix} \qquad (6.1)$$

where $y \in R^{4k+m}$, y_i unrestricted for $i=1,\ldots,k$, and $y_i \geq 0$ for $i=k+1,\ldots,4k+m$. Hence, an elimination of some of the column vectors in the model is possible. However, the essence of the computation lies in the number of constraints since this determines the size of the bases which must be manipulated at each iteration of the simplex algorithm. A reduction in the size of the working basis is possible in this problem. To demonstrate this reduction, let us express the last k constraints on the model as

$$\begin{bmatrix} I_k & I_k & I_k \end{bmatrix} \bar{y} = \begin{pmatrix} p \\ \vdots \\ p \end{pmatrix}_k \qquad (6.2)$$

$\bar{y} \varepsilon R^{3k}$, $\bar{y} \geq 0$. Note that the system may be expressed as

$$\begin{bmatrix} I_k & I_k \end{bmatrix} y^* \leq \begin{pmatrix} p \\ \vdots \\ p \end{pmatrix} \qquad (6.3)$$

$$y^* \varepsilon R^{2k},\ y^* \geq 0$$

since the last k components of the vector $\bar{y}$ are, effectively, "slack" variables. Moreover, rearranging the order of the variables in (6.3) suggests that we consider the equivalent system:

$$\begin{bmatrix} 1 & 1 & & & & & \\ & & 1 & 1 & & & \\ & & & & \ddots & 1 & 1 \end{bmatrix} \bar{y}^* \leq \begin{pmatrix} p \\ \vdots \\ p \end{pmatrix}_k .$$

Observe that this set of constraints is amenable to the methods of Generalized Upper Bounding. In particular, this formulation permits the last k constraints in the linear system (6.1) to be removed from the working basis when this upper bounding algorithm is employed.

In summary, the constraint set for the dual problem, Problem Q, contains $5k+m$ variables and $2k+n$ constraints with a working basis of rank $2k+n$. The results of this analysis indicate that this dual problem can be expressed equivalently as a linear system involving $4k+m$ variables, $k+n$ "natural" constraints, and a system of k "generalized upper bounding" constraints. Thus, the effective working basis has rank $k+n$. This reduction could be very significant in the solution of large scale problems--particularly when there are many goal functions in the model. Clearly, further study of this dual problem, Problem Q, might yield further reductions to expedite the solution of large models.

6.2 Concluding Remarks

Perhaps the most significant contribution of this research is the potential application of these results to multiple criterion optimization where the decision-maker is interested in studying alternative measures of achievement based on the ℓ_p metric. Currently, such analysis requires the availability of nonlinear programming software. Although these results demonstrate that linear optimization techniques are sufficient, it is clear that the size of the resulting linear model may become a deterrent for large-scale problems. In summary, it appears that more and larger problems were uncovered than solved.

7. BIBLIOGRAPHY

1. G. Appa and C. Smith, "On L_1 and Chebyshev Estimation," Mathematical Programming, 5, 73-87 (August 1973).

2. I. Barrodale and F. Roberts, "Applications of Mathematical Programming to ℓp Approximation," in J. Rosen, O. Mangasarian, K. Ritter (eds.) Nonlinear Programming, Academic Press, New York, 1970.

3. R. Benayoun, J. de Montogolfier, and O. Laritchev, "Linear Programming with Multiple Objective Functions: Step Method (STEM)," Mathematical Programming, 1, 366-375 (December 1971).

4. R. Benayoun, J. Tergny, and D. Keuneman, "Mathematical Programming with Multi-Objective Functions: A Solution by P.O.P. (Progressive Orientation Procedure)," Metra, 9, 279-299 (June 1970).

5. R. Brandis, "Principles of Economics," Richard D. Irwin, Inc., Homewood, Illinois, 1972.

6. A. Charnes and W. W. Cooper, "Goal Programming and Constrained Regression - A Comment," OMEGA, 3, 403-309 (April 1975).

7. A. Charnes and W. W. Cooper, "Management Models and Industrial Applications of Linear Programming Volume I," John Wiley and Sons, Inc., New York, 1961.

8. A Charnes, W. W. Cooper, D. Klingman, and R. J. Niehaus, "Explicit Solutions to Convex Goal Programming," Management Science, 22, 438-448 (December 1975).

9. A Charnes, W. W. Cooper, R. J. Niehaus, and D. Scholtz, "A Multiple-Objective Model for Planning Equal Employment Opportunities," in M. Zeleny (ed.) Multiple Criterion Decision Making: Kyoto 1975, Springer-Verlag, New York, 1975.

10. A. Charnes, W. W. Cooper, and Y. Ijiri, "Breakeven Budgeting and Programming Goals," Journal of Accounting Research, 1, 16-43 (January 1963).

11. N. V. Chernikova, "Algorithm for Finding a General Formula for the Nonnegative Solutions of a System of Linear Equations," USSR Computational Mathematics and Mathematical Physics, 4, (4), 151-158 (1964).

12. H. Chernoff and L. E. Moses, "Elementary Decision Theory," John Wiley and Sons, Inc., New York, 1959.

13. J. L. Cohen and D. H. Marks, "A Review and Evaluation of Multiobjective Programming Techniques," Water Resources Research, 11, 208-220 (April 1975).

14. B. Contini, "A Stochastic Approach to Goal Programming," Operations Research, 16, 578-586 (May-June 1968).

15. N. R. Draper and H. Smith, "Applied Regression Analysis," John Wiley and Sons, Inc., New York, 1966.

16. R. J. Duffin, E. L. Peterson, and C. Zener, "Geometric Programming - Theory and Applications," John Wiley and Sons, Inc., New York, 1967.

17. J. Dyer, "Interactive Goal Programming," Management Science, 19, 62-70 (September 1972).

18. J. Ecker and I. Kovada, "Finding Efficient Points for Linear Multiple Objective Programs," Mathematical Programming, 8, 375-377 (June 1975).

19. J. P. Evans and R. E. Steuer, "A Revised Simplex Method for Linear Multiple Objective Programs," Mathematical Programming, 5, 54-72 (August 1973).

20. J. P. Evans and R. E. Steuer, "Generating Efficient Extreme Points in Linear Multiple Objective Programming: Two Algorithms and Computing Experience," in J. L. Cochrane and M. Zeleny (eds.) Multiple Criterion Decision Making, University of South Carolina Press, Columbia, South Carolina, 1973.

21. P. C. Fishburn, "Utility Theory for Decision Making," John Wiley and Sons, Inc., New York, 1970.

22. A. B. Forsythe, "Robust Estimation of Straight Line Regression Coefficients by Minimizing pth Power Deviations," Technometrics, 14, 159-166 (February 1972).

23. A. Geoffrion, "Proper Efficiency and the Theory of Vector Maximization," Journal of Mathematical Analysis and Applications, 22, 618-620 (June 1968).

24. A. Geoffrion, J. S. Dyer, and A. Feinberg, "An Interactive Approach for Multi-Criterion Optimization with an Application to the Operation of an Academic Department," Management Science, 19, 357-368 (December 1972).

25. G. Hadley, "Introduction to Probability and Statistical Decision Theory," Holden-Day, Inc., San Francisco, California, 1967.

26. Y. Ijiri, "Management Goals and Accounting for Control," North-Holland Publishing Company, Amsterdam, 1965.

27. S. Karlin, "Mathematical Methods and Theory of Games, Programming, and Economics," Addison-Wesley Publishing Company, Inc., Reading, Massachusetts, 1959.

28. C. N. Klahr, "Multiple Objectives in Mathematical Programming," Operations Research, 6, 849-855 (November-December 1958).

29. T. C. Koopmans, "Analysis of Production as an Efficient Combination of Activities," in Activity Analysis of Production and Allocation, Cowles Commission Monograph No. 13, John Wiley and Sons, Inc., New York, 1951.

30. J. Kornbluth, "A Survey of Goal Programming," OMEGA, 1, 193-205 (April 1973).

31. J. Kornbluth, "Duality, Indifference and Sensitivity Analysis in Multiple Objective Linear Programming," Operational Research Quarterly, 25, 599-614 (December 1974).

32. O. Krafft, "Programming Methods in Statistics and Probability Theory," in J. Rosen, O Mangasarian, K. Ritter (eds.) Nonlinear Programming, Academic Press, New York, 1970.

33. H. W. Kuhn and A. W. Tucker, "Nonlinear Programming," in Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability, University of California Press, Berkeley, California, 1950.

34. S. Lee, "Goal Programming for Decision Analysis, "Auerbach Publishers, Inc., Philadelphia, Pennsylvania, 1972.

35. R. Lehmann and W. Oettli, "The Theorem of the Alternative, the Key Theorem, and the Vector-Maximum Problem," Mathematical Programming, 8, 332-344 (June 1975).

36. D. G. Luenberger, "Optimization by Vector Space Methods," John Wiley and Sons, Inc., New York, 1969.

37. O. Mangasarian, "Nonlinear Programming," McGraw-Hill Book Company, New York, 1969.

38. M. Markowitz, "The Optimization of a Quadratic Function Subject to Linear Constraints," _Naval Research Logistics Quarterly_, _3_, 14-27 (January 1956).

39. J. Philip, "Algorithms for the Vector Maximum Problem," _Mathematical Programming_, 2, 207-229 (April 1972).

40. B. Roy, "Problems and Methods with Multiple Objective Functions," _Mathematical Programming_, _1_, 239-266 (November 1971).

41. V. A. Sposito, "Linear and Nonlinear Programming," Iowa State University Press, Ames, Iowa, 1975.

42. R. E. Steuer, "Linear Multiple Objective Programming: Theory and Computation," unpublished Ph.D. thesis, University of North Carolina (1973).

43. J. Von Neumann and O. Mortgenstern, "Theory of Games and Economic Behavior," Princeton University Press, Princeton, New Jersey, 1944.

44. H. M. Wagner, "Linear Programming Techniques for Regression Analysis," _Journal of American Statistical Association_, _54_, 206-212 (March 1959).

8. ACKNOWLEDGEMENTS

A sincere debt of appreciation is expressed to Howard Meeks and Vince Sposito for their encouragement and support throughout my research and preparation of this thesis. Moreover, their efforts on my behalf have again established the fact that success in graduate studies is a function of motivation and perserverance and not a function of innate intelligence.

In retrospect, it is clear that the most significant contributors to my overall graduate education were the many students it has been my privilege and pleasure to teach. Collectively these individuals have contributed more to my professional growth and development than I know how to acknowledge. Accordingly, a sincere debt of gratitude is expressed to Keith McRoberts for providing me with the opportunity to teach and hence learn and grow at a rate that surpassed my expectations.

9. APPENDIX A: GENERALIZED INVERSE OF A MATRIX

Let C be any arbitrary matrix. Then the generalized inverse of C, denoted as C^*, may be defined as that unique matrix which satisfies the following four equations:

1. $CC^*C = C$
2. $C^*CC^* = C^*$
3. $CC^* = (CC^*)'$
4. $C^*C = (C^*C)'$

It can be shown that for any matrix C (nonsingular, singular, square, rectangular, zero or nonzero) there exists a unique matrix C^* which satisfies the above set of equations.

When a matrix C has an ordinary inverse C^{-1} (i.e., when C is nonsingular), C^* is equivalent to C^{-1} since C^{-1} satisfies the first equation, and the uniqueness property of C^* guarantees that $C^* = C^{-1}$. As shown in Ijiri [26], the generalized inverse C^* of a matrix C has the following properties:

1. $C = 0$ (mxn) implies that $C^* = 0$ (nxm).
2. $C^{**} = C$.
3. $(C')^* = (C^*)'$.
4. If C is nonsingular, then $C^* = C^{-1}$.
5. $(C'C)^* = C^*(C^*)'$.
6. If U and V are unitary, then $(UCV)' = V'C^*U'$.
7. If $C = \Sigma A_i$ where $C_iC_j' = 0$ and $C_i'C_j = 0$ whenever $i \neq j$, then $C^* = \Sigma C_i^*$.

8. If C is normal, then $C^*C = CC^*$ and $(C^n)^* = (C^*)^n$.

9. $C, C'C, C'$, and C^*C all have rank equal to trace C^*C.

10. C^*C, CC^*, $(I - C^*C)$, and $(I - CC^*)$ are all hermitian and idempotent.

11. $(aC)^* = a^*C^*$ where a is a complex number and a^* means a^{-1} if $a \neq 0$ and 0 if $a = 0$.

12. If C is hermitian and idempotent, then $C^* = C$.

13. If C has full column rank, then $C^* = (C'C)^{-1}C'$.

14. If C has full row rank, then $C^* = C'(CC')^{-1}$.

15. If $B(m \times r)$, $C(r \times r)$, and $D(r \times n)$ each has rank r, where $1 \leq r \leq$ minimum (m,n), then $(BCD)^* = D^*C^*B^*$.

10. APPENDIX B: DUALITY IN GEOMETRIC PROGRAMMING

In this appendix we will consider the fundamental properties of primal and dual geometric programming problems. In particular, the duality relationship itself will be explored as the dual problem is constructed. Virtually all of the material presented herein may be found in Duffin, Peterson, and Zener [16]; hence, explicit references on the key theorems will be omitted.

The most general form of a primal geometric programming problem is given as:

PROBLEM GP

$$\text{minimize} \quad g_o(x)$$

subject to

$$g_k(x) \leq 1 \qquad k = 1,2,\ldots,p$$

$$x_j > 0 \qquad j = 1,2,\ldots,n$$

Here

$$g_k(x) = \sum_{i=1}^{m_k} c_{ki} \prod_{j=1}^{n} x_j^{a_{kij}} \qquad k = 0,1,\ldots,p \quad (B.1)$$

where $c_{ki} > 0$ and a_{kij} denote arbitrary real numbers.

To obtain notational simplicity we will express (B.1) as

$$g_k(x) = \sum_{i \varepsilon I[k]} c_i \prod_{j=1}^{n} x_j^{a_{ij}} \qquad k = 0,1,\ldots,p$$

where I [k] denotes the appropriate set of indices for the function $g_k(x)$ such that $I[0] = \{1,2,\ldots,m_o\}$, $I[1] = \{m_o+1, \ldots, m_o+m_1\}$ etc.

The functions $g_k(x)$ are known as posynomials since each term is guaranteed to be positive over its domain of definition. As a consequence of these positive terms, geometric programming is a branch of convex programming. However, the highly nonlinear nature of this programming problem suggests that a solution procedure should be based on an equivalent problem which is more computationally attractive. Such is the nature of the dual geometric programming problem. Before presenting the dual problem, we will introduce several key results which will be useful in the analysis of the duality relationship.

Duality in geometric programming is based on an application of the well-known arithmetic mean-geometric mean inequality. In its full generality, this inequality may be stated as follows: If $\mu_1, \mu_2, \ldots \mu_n$ are n nonnegative numbers and if $\delta_1, \delta_2, \ldots \delta_n$ have the property that

$$\sum_{i=1}^{n} \delta_i = 1$$

and

$$\delta_i > 0 \qquad \text{for} \qquad i = 1,\ldots,n,$$

then

$$\sum_{i=1}^{n} \delta_i \mu_i \geq \prod_{i=1}^{n} \mu_i^{\delta_i} \tag{B.2}$$

In the special case where $\delta_i = 1/n$, the left-hand side of (B.2) and the right-hand side of (B.2) are, by definition, the arithmetic mean and the geometric mean, respectively. Hence, in this case, (B.2) may be expressed as

$$\frac{1}{n} \sum_{i=1}^{n} \mu_i \geq \prod_{i=1}^{n} \mu_i^{1/n}$$

which is the familiar arithmetic mean--geometric mean inequality. An extension of this classic inequality is now presented for future reference in the following lemma.

<u>LEMMA B.1</u> Let $\mu_i > 0$, $\delta_i \geq 0$ for $i=1,\ldots,n$ be arbitrary real numbers. Then

$$\left(\sum_{i=1}^{n} \mu_i\right)^{\lambda} \geq \prod_{i=1}^{n} \left(\frac{\mu_i}{\delta_i}\right)^{\delta_i} \lambda^{\lambda}$$

where

$$\lambda = \sum_{i=1}^{n} \delta_i$$

and

$$\left(\frac{\mu_i}{\delta_i}\right)^{\delta_i} = 1$$

if $\delta_i = 0$. Moreover, the inequality becomes an inequality if, and only if,

$$\delta_j \sum_{i=1}^{n} \mu_i = \mu_j \sum_{i=1}^{n} \delta_i, \qquad j=1,\ldots,n.$$

The role of this inequality is central to the theory of duality in that it provides a basis for the proof of the Main Lemma of Geometric Programming and, hence, the weak and strong duality theorems.

In developing the most general form of the dual geometric program we first consider the construction of a dual objective function tailored to the unconstrained minimization of a posynomial

$$g(x) = \sum_{j=1}^{m} \mu_1.$$

An application of Lemma B.1 states that

$$\delta_i U_1 + \ldots + \delta_m U_m \geq U_1^{\delta_1} \ldots U_m^{\delta_m} \tag{B.3}$$

where $U_1,\ldots,U_m$ are arbitrary nonnegative numbers and $\delta_1,\ldots\delta_m$ are positive weights which satisfy the normality condition

$$\sum_{i=1}^{m} \delta_i = 1$$

Letting $\mu_i = \delta_i U_i$ implies that

$$\sum_{i=1}^{m} \mu_i \geq \prod_{i=1}^{m} \left(\frac{\mu_i}{\delta_i}\right)^{\delta_i} \tag{B.4}$$

Substituting the terms $\mu_i = c_i \prod_{j=1}^{n} x_j^{a_{ij}}$ into the right side of (B.4) we have the pre-dual function

$$v(\delta,x) = \prod_{i=1}^{m} \left(\frac{c_i \prod_{j=1}^{n} x_j^{a_{ij}}}{\delta_i}\right)^{\delta_i}$$

$$= \prod_{i=1}^{m} \left(\frac{c_i}{\delta_1}\right)^{\delta_i} \prod_{j=1}^{n} x_j^{S_j}$$

where S_j denotes the linear combinations

$$S_j = \sum_{i=1}^{m} a_{ij}\delta_i \qquad \text{for} \qquad j=1,\ldots,n.$$

Suppose now, that it is possible to select the weights, δ_i, such that $S_j = 0$ for all j. Then the pre-dual function is independent of the primal variables x_j. That is, we restrict the dual variables to be contained in a dual space Ω_D which is the orthogonal complement of the primal space Ω_P. The result, then, is the dual function

$$v(\delta) = \prod_{i=1}^{m} \left(\frac{c_i}{\delta_i}\right)^{\delta_i} .$$

Note also, in view of (B.4) that

$$g(x) \geq M \geq v(\delta)$$

for any $x \in \Omega_P$, $\delta \in \Omega_D$.

Incorporating a prototype posynomial constraint into the program introduces a set of unnormalized weights. Let $\Delta_1, \ldots, \Delta_n$ denote these unnormalized weights and by $\lambda(\Delta)$ the sum:

$$\lambda(\Delta) = \sum_{i=1}^{n} \Delta_i .$$

The relationship between the normalized weights δ and the unnormalized weights Δ is then $\Delta_i = \lambda(\Delta)\delta_i$ which implies that

$$\delta_i = \frac{\Delta_i}{\lambda(\Delta)} \qquad \text{for} \qquad i = 1,\ldots,m \qquad \text{(B.5)}$$

For unnormalized weights, we have as a consequence of Lemma B.1 that

$$\sum_{i=1}^{m} \mu_i \geq \prod_{i=1}^{m} \left[\left(\frac{\mu_i}{\Delta_i} \right)^{\Delta_i/\lambda} \right] \lambda(\Delta).$$

In view of the above we can state a prototype constrained geometric program as follows

$$\text{MIN } g_o^{\lambda_o(\delta)} \geq \prod_{i=1}^{n} \left[\left(\frac{\mu_i}{\Delta_i} \right)^{\Delta_i} \right] \lambda_o(\Delta)^{\lambda_o(\Delta)} \qquad \text{(B.6)}$$

subject to

$$1 \geq g(x)^{\lambda(\Delta)} \geq \prod_{i=n+1}^{M} \left[\left(\frac{\mu_i}{\Delta_i} \right)^{\Delta_i} \right] \lambda(\Delta)^{\lambda(\Delta)} \qquad \text{(B.7)}$$

Now, multiplying inequality (B.6) by (B.7) we have that

$$g_o(x)^{\lambda_o(\Delta)} \geq \prod_{i=1}^{m} \left[\left(\frac{\mu_i}{\Delta_i} \right)^{\Delta_i} \right] \prod_{i=m+1}^{M} \left[\left(\frac{\mu_i}{\Delta_i} \right)^{\Delta_i} \right] \lambda_o(\Delta)^{\lambda_o(\Delta)} \; \lambda(\Delta)^{\lambda(\Delta)}.$$

This inequality is valid for any selection of Δ. It is more meaningful, however, to select the normalization $\lambda_o(\Delta) = 1$. Letting δ_i denote the weights normalized in this manner we have that

$$g_o(x) \geq \prod_{i=1}^{m} \left[\left(\frac{\mu_i}{\delta_i} \right)^{\delta_i} \right] \lambda(\delta)^{\lambda(\delta)} = V(\delta, x)$$

$$= \prod_{i=1}^{m} \left[\left(\frac{c_i}{\delta_i} \right)^{\delta_i} \right] \lambda(\delta)^{\lambda(\delta)} \prod_{j=1}^{n} x_j^{\sum_{i=1}^{m} a_{ij}\delta_i} = v(\delta) \prod_{j=1}^{n} x_j^{S_j}.$$

Likewise, this procedure can be extended to a program with N terms and p prototype constraints. In this case,

$$v(\delta) \prod_{j=1}^{n} x_j^{S_j} = \prod_{i=1}^{N} \left[\left(\frac{c_i}{\delta_i} \right)^{\delta_i} \right] \prod_{k=1}^{p} \lambda_i(\delta)^{\lambda_i(\delta)} \prod_{j=1}^{n} x_j^{S_j} .$$

Again, restricting δ_i to the dual space Ω_D forces S_j to vanish for all j and gives the desired result. In the constrained case, as well as the unconstrained case, we also have that

$$g_o(x) \geq M \geq v(\delta)$$

for any

$$x \in \Omega_P, \ \delta \in \Omega_D.$$

Thus, the dual of Problem GP is given as

<u>PROBLEM GD</u>

$$\text{Maximize} \quad v(\delta) = \prod_{i=0}^{p} \prod_{i \in I[k]} \left[\left(\frac{c_i}{\delta_i} \right)^{\delta_i} \right] \prod_{k=1}^{p} \lambda_k(\delta)^{\lambda_k(\delta)}$$

subject to

$$\lambda_o(\delta) = \sum_{i \in I[o]} \delta_i = 1 \qquad \text{(normality)}$$

$$\sum_{i=1}^{N} a_{ij}\delta_i = 0 \quad j=1,\ldots,n \qquad \text{(orthogonality)}$$

$$\delta_i \geq 0 \qquad i = 1,\ldots,N \quad \text{(positivity)}$$

Consider, now, the relationship between the primal program, Problem GP, and the dual program, Problem GD. This duality relationship is characterized by the <u>Main Lemma of Geometric Programming</u> which is given as

<u>LEMMA B.2</u> If x satisfies the constraints of the primal problem and δ satisfies the constraints of the dual problem, then

$$g_o(x) \geq v(\delta).$$

Moreover, under the same conditions,

$$g_o(x) = v(\delta)$$

if, and only if,

$$\delta_i = \begin{cases} \dfrac{c_i \prod_{j=1}^{n} x_j^{a_{ij}}}{g_o(x)} & i \varepsilon I\,[0] \\ \lambda_k(\delta)\; c_i \prod_{j=1}^{n} x_j^{a_{ij}} & i \varepsilon I\,[k],\ k=1,\ldots,p. \end{cases}$$

With regard to the equivalence of the primal and dual problems, necessity and sufficiency are provided by the first and second duality theorems of geometric programming which are given as:

THEOREM B.3 Suppose Problem GP is superconsistent (i.e., satisfies Slater's condition) and that the primal function $g_o(x)$ attains its minimum value at a point which satisfies the primal constraints. Then

1. The corresponding dual program, Problem GD, is consistent and the dual function $v(\delta)$ attains its constrained maximum at a point which satisfies the dual constraints.

2. The constrained maximum value of the dual function is equal to the constrained minimum value of the primal function.

3. If x^* is a minimizing point for Problem GP, then there are nonnegative Lagrangian multiplious y_k^*, $k=1,\ldots,p$, such that the Lagrangian function

$$L(x,y) = g_o(x) + \sum_{k=1}^{p} y_k(g_k(x) - 1)$$

has the property

$$L(x^*,y) \le g_o(x^*) = L(x^*,y^*) \le L(x,y^*)$$

for arbitrary $x_j > 0$ and arbitrary $y_k \geq 0$. Moreover, there is a maximizing vector δ^* such that

$$\delta_i^* = \begin{cases} \dfrac{c_i \prod_{j=1}^{n} x_j^{a_{ij}}}{g_o(x)}, & i \epsilon I\ [0] \\[2ex] \dfrac{y_k c_i \prod_{j=1}^{n} x_j^{a_{ij}}}{g_o(x)}, & i \epsilon I\ [k], \qquad k=1,\ldots,p \end{cases}$$

where $x = x^*$ and $y = y^*$. Furthermore,

$$\lambda_k(\delta) = \frac{y_k^*}{g_o(x^*)}, \qquad k=1,2,\ldots,p$$

4. If δ^* is a maximizing point for dual Problem GD, each minimizing point x^* for primal Problem GP satisfies the system

$$c_i \prod_{j=1}^{n} x_j^{a_{ij}} = \begin{cases} \delta_i^* \, v(\delta^*), & i \epsilon I\ [0] \\ \delta_i^*/\lambda_k(\delta^*), & i \epsilon I\ [k] \end{cases}$$

where k ranges over all positive integers for which $\lambda_i(\delta^*) > 0$.

<u>THEOREM B.4</u> If primal Problem GP is consistent and there is a point δ^* with positive components which satisfies the constraints

of dual Problem GD, the primal function $g_o(x)$ attains its constrained minimum value at a point x^* which satisfies the constraints of primal Problem GP.

11. APPENDIX C: THE RELATIONSHIP BETWEEN LINEAR AND GEOMETRIC PROGRAMMING

On studying the properties of posynomial functions on geometric programming, Alex Federowicz in [16] observed a peculiar relationship between linear programming and geometric programming. Namely, by a simple transformation of variables, the equivalence between linear programming and geometric programming with single-term functions is easily established. His analysis follows.

PROBLEM L1

$$\text{Minimize} \quad G_o(z) = a_{01}z_1 + a_{02}z_2 + \dots + a_{0n}z_n + C_0$$

subject to

$$G_i(z) = a_{i1}z_1 + a_{i2}z_2 + \dots + a_{in}z_n + C_i \leq 0, \quad i = 1,\dots,m$$

where a_{ij} and C_i denote arbitrary constants. Using the following one-to-one transformations:

$$G_i(z) = \ell n\ g_i,\ C_i = \ell n\ c_i,\ z_j = \ell n\ x_j,$$

where each c_i and x_j is positive, we can express Problem L1 as an equivalent geometric program.

PROBLEM M1

Minimize $$g_0(x) = c_0 \prod_{j=1}^{n} x_j^{a_{0j}}$$

subject to

$$x_j > 0 \qquad j = 1,\ldots,n$$

and

$$g_i(x) = c_i \prod_{j=1}^{n} x_j^{a_{ij}} \leq 1 \qquad i = 1,\ldots,m.$$

Note that this program is a special type of geometric program for which there is only one term in each posynomial; such a single-term posynomial is called a monomial. Accordingly, the dual problem is expressed as:

PROBLEM M2

Maximize $$v(\delta) = \left[\prod_{i=0}^{m} \left(\frac{c_i}{\delta_i} \right)^{\delta_i} \right] \prod_{i=1}^{m} \lambda_i(\delta)^{\lambda_i(\delta)} \qquad \text{(C.1)}$$

subject to

$$\delta_i \geq 0 \qquad i=1,\ldots,m,$$

$$\delta_0 = 1.$$

and

$$\sum_{i=1}^{m} a_{ij} \delta_i = 0 \qquad j=1,\ldots,n.$$

Here a_{ij} and c_i are the exponents and coefficients, respectively, as given in Problem M1. Federowicz also observed that this dual program can be further simplified by exploiting the monomial form of its primal. Restating the product function (C.1) we have that

$$v(\delta) = \left[\prod_{i=0}^{m} \left(\frac{c_i}{\delta_i}\right)^{\delta_i}\right] \prod_{i=1}^{m} \lambda_i(\delta)^{\lambda_i(\delta)}$$

where the dual dependent variable $\lambda_i(\delta)$ is defined as

$$\lambda_i(\delta) = \sum_{i \in J[i]} \delta_i .$$

Note that, for a program with monomial constraints,

$$\lambda_i(\delta) = \delta_i .$$

Thus,

$$v(\delta) = \prod_{i=o}^{m} \left(\frac{c_i}{\delta_i}\right)^{\delta_i} \prod_{i=1}^{m} \delta_i^{\ \delta_i}$$

$$= \left(\frac{c_o}{\delta_o}\right)^{\delta_o} \prod_{i=1}^{m} \left(\frac{c_i}{\delta_i}\right)^{\delta_i} \prod_{i=1}^{m} \delta_i^{\delta_i}$$

$$= c_o \prod_{i=1}^{m} \left[\left(\frac{c_i}{\delta_i}\right)^{\delta_i} (\delta_i)^{\delta_i} \right]$$

$$= c_o \prod_{i=1}^{m} \left[(c_i)^{\delta_i} \ (\delta_i)^{-\delta_i} \ (\delta_i)^{\delta_i} \right]$$

$$= c_o \prod_{i=1}^{m} (c_i)^{\delta_i} .$$

Now, using the one-to-one transformation

$$V(\delta) = \ln [v(\delta)]$$

we can express Problem M2 equivalently as:

PROBLEM M3

Maximize $V(\delta) = C_o + \sum_{i=1}^{m} C_i \delta_i$

where $C_i = \ln(c_i)$ for $i = 0,1,\ldots,m,$

subject to $\delta_i \geq 0,$ $i = 0,1,\ldots,m,$

and

$$a_{o1} + a_{11}\delta_1 + \cdots + a_{m1}\delta_m = 0$$

$$a_{o2} + a_{12}\delta_2 + \cdots + a_{m2}\delta_m = 0$$

$$\vdots \quad \vdots \quad \vdots \quad \vdots$$

$$a_{on} + a_{1n}\delta_1 + \cdots + a_{mn}\delta_m = 0$$

Observe that the resulting linear program is simply the dual of Problem L1 which is as it should be.