AD-A059 859 CALIFORNIA UNIV BERKELEY OPERATIONS RESEARCH CENTER F/G 12/1

LINEAR DECISION MODELS UNDER RISK.(U)

APR 78 T R HARMS N00014-76-C-0134

UNCLASSIFIED ORC-78-7 NL

1 OF 2

AD
A059859

(12) LEVEL III

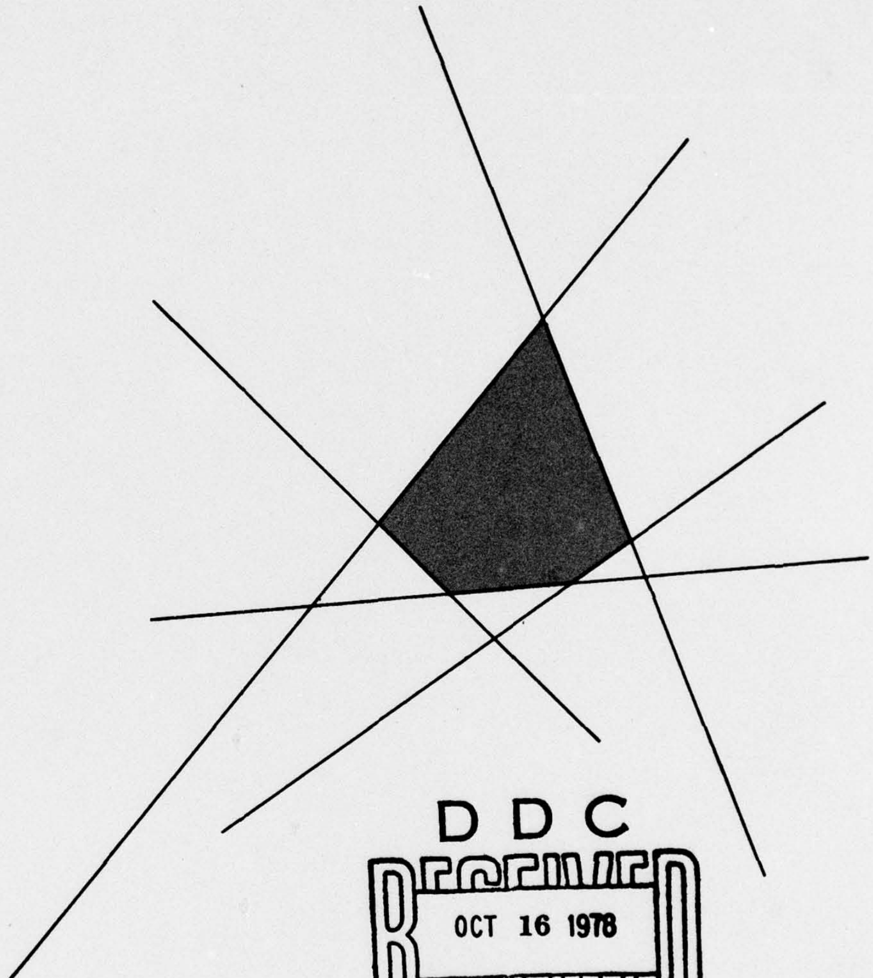
NW

ORC 78-7
APRIL 1978

LINEAR DECISION MODELS UNDER RISK

by

TERRY R. HARMS



OPERATIONS
RESEARCH
CENTER

UNIVERSITY OF CALIFORNIA · BERKELEY

6

LINEAR DECISION MODELS UNDER RISK

by

10 Terry R. Harms
Operations Research Center
University of California, Berkeley

9 Research rept.,

15 N00014-76-C-0134

12 131p.

11 APRIL 1978

14 ORC-78-7

This research was supported by the Office of Naval Research under Contract N00014-76-C-0134 with the University of California. Reproduction in whole or in part is permitted for any purpose of the United States Government.

270 750

78 10 10 123

mt

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: ORC 78-7
2. GOVT ACCESSION NO.:
3. RECIPIENT'S CATALOG NUMBER:
4. TITLE (and Subtitle): LINEAR DECISION MODELS UNDER RISK
5. TYPE OF REPORT & PERIOD COVERED: Research Report
6. PERFORMING ORG. REPORT NUMBER:
7. AUTHOR(s): Terry R. Harms
8. CONTRACT OR GRANT NUMBER(s): N00014-76-C-0134
9. PERFORMING ORGANIZATION NAME AND ADDRESS: Operations Research Center, University of California, Berkeley, California 94720
10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: NR 047 033
11. CONTROLLING OFFICE NAME AND ADDRESS: Office of Naval Research, Department of the Navy, Arlington, Virginia 22217
12. REPORT DATE: April 1978
13. NUMBER OF PAGES: 130
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office):
15. SECURITY CLASS. (of this report): Unclassified
15a. DECLASSIFICATION/DOWNGRADING SCHEDULE:
16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Stochastic Programming
Chance Constraints

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

(SEE ABSTRACT)

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

S/N 0102-LF-014-6601

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

DEDICATION

To Celia

ACKNOWLEDGEMENTS

I would like to express my appreciation to my thesis advisor, Professor Ilan Adler, for the many suggestions and the encouragement he provided me. I also wish to thank Professors C. Roger Glassey and Pravin Varaiya for serving on my dissertation committee, and the entire faculty of the Department of Industrial Engineering and Operations Research at the University of California, Berkeley, in particular, Professor R. W. Shephard, for helping me throughout my graduate studies at Berkeley. I would also like to thank the entire Operations Research Center staff, especially Mrs. Gloria Partee for her excellent typing.

Finally, I thank my wife Celia, not only for her understanding and patience, but also for her support throughout this undertaking.

ABSTRACT

The assumption made in linear programming that the components are deterministic (constant) numbers is rarely fulfilled in practical applications. This has led to the development of the field of stochastic programming where the random aspect of the coefficients in the objective function, technology matrix, and the vector of resources are taken into account. ~~In~~ this research ~~we~~ investigates the problem of a linear program with uncertainty attached to the decision vector. For example, a decision to order a certain amount of a perishable good might yield variable amounts of this good at delivery due to spoilage.

Two models are considered:

1. The uncertainty is independent of the decision. A decision x will yield an output $x + e$ where e is a random variable.
2. The uncertainty is proportional to the decision. A decision x will yield an output $x + \alpha x$ where α is a random variable.

As can be seen in the literature on stochastic programming, the random nature of the program does not lead to a unique mathematical problem and there are various models of stochastic linear programs. In this study we chose to use the chance-constrained approach to formulate two models incorporating the two kinds of decision uncertainties described above. In chance-constrained programs the criterion is the expected value of the

objective function and the constraints have to be satisfied within a predetermined fixed probability.

Using known methods of chance-constrained programming we define deterministic equivalent problems for our two stochastic models. It is assumed that the distributions of the random variables are known and convolutions of these distributions can easily be obtained. In general this is not the case, this is why we then proceed to find conservative approximations for our two models (i.e., problems with the same objective functions whose feasibility sets are subsets of the original feasibility sets). We present a set of random variables for which conservative approximations are easily obtainable linear programs. Bounds on the optimal value of the objective function are also defined. Using these results, different kinds of sensitivity analyses are investigated. In this context a problem of trade-off between the cost of reducing the variance of the uncertainties versus the corresponding improvement in the objective function is defined and a simple algorithm is presented to solve this problem.

Finally, we outline how some of the results can be applied to more general chance-constrained programs and conservative approximations can be defined for these more general models.

TABLE OF CONTENTS

INTRODUCTION

Stochastic Linear Programming is the field of study of linear programming problems, where the random aspect of the coefficients in the objective function, technology matrix, and the vector of resources are taken into account. More than one approach has been developed in this context such as the distribution problem [13] and multistage Recourse Models [13]. In this research we will restrict our study to the chance-constrained approach, a method originally introduced by Charnes and Cooper [3]. Chance-constrained programming allows constraint violations up to specified probability limits. This study differs from the usual chance-constrained problem in that the randomness in the decision vectors is studied rather than the coefficients of the objective function, technology matrix, and vector of resources. For example, a decision to order a certain amount of a perishable good might yield variable amounts of this good at delivery due to spoilage. Another example is the uncertainty due to measuring errors. These are the types of uncertainty that we are trying to incorporate into a linear program.

In Chapter I we present two basic models which will be analyzed throughout this study. In Model I, the uncertainty is independent of the decision. A decision x will yield an output $x + e$ where e is a random variable. In Model II, the uncertainty is proportional to the decision. A decision x will yield an output $x + \alpha x$ where α is a random variable. Using the approach developed by Charnes and Cooper [4], deterministics equivalents for Models I and II are presented and discussed. Because of the difficulty of implementing in practice these deterministic equivalents, we looked for a class of distributions for which approximations could be determined with little effort. To this effect, in Chapter II a class of densities and the corresponding set of random variables and their properties are presented. In Chapter III, the results of Chapter II are exploited to find "conservative approximations" to Models I and II under certain assumptions. By "conservative approximation" we mean a problem with the same objective function as the original problem and a feasibility set which is a subset of the original feasibility set. In Chapter IV, sensitivity analysis of Models I and II are discussed with respect to the cost vector, the vector of resources, the specified probability limits, and variations in the densities of the random variables. Finally, the "reducing of uncertainty" problem is presented. It is a problem of trade-off between the cost of decreasing the uncertainty by modifying the densities of the random variables versus the original linear objective function.

In Chapter V an algorithm is presented for solving a special kind of reducing of uncertainty problem which has the following general formulation.

$$\begin{aligned} \text{minimize} \quad & c^T x + \frac{1}{d^T y} \quad (\text{with } d > 0) \\ \text{subject to} \quad & Ax - By \geq b \\ & x \geq 0 \\ & 1 \geq y \geq 0 \end{aligned}$$

where A , B , b , c , d are given, A and B are $m \times n$ matrices, c and d are $n \times 1$ vectors and b is an $m \times 1$ vector. x and y

are $n \times 1$ vectors of decisions.

Finally in Chapter VI, the basic two models are generalized to joint-chance constraint models as developed in Miller and Wagner [10]. The problems here are different from the usual joint-chance constrained programming in that the constraints are not statistically independent as is usually assumed. This is why we use the concept of associativity of random variables, widely used in the field of reliability [2], and the results of Chapter II to define conservative approximations for these generalized models.

CHAPTER I

A. Introduction

This research is concerned with a special kind of stochastic program. We are interested in the usual linear program with one important added feature. The decisions have uncertainty attached to them. For example, a decision to order a certain amount of a perishable good might in reality yield variable amounts at delivery. Another example is the uncertainty due to errors of the measuring instruments. It is this type of uncertainty that we are trying to incorporate in a regular linear program. Of course the stochastic nature of the uncertainties changes the character of the problem. In this research we will use a chance constrained approach to the problem. The linear constraints are to be satisfied within a certain specified probability. The objective function will be the usual expected value criterion. Two models will be presented.

B. Model I

First we will look at the following approach. Each variable x_j has an uncertainty attached to it in the form of the random variable e_j . The a priori decision is x_j and the outcome is $x_j + e_j$. This model fits well in the context of measurement errors. Within certain bounds the error of measurement is a function of the measuring instrument, not the quantity measured and therefore the error e_j and the quantity x_j are independent. This yields the following model:

Model I:

$$\text{minimize} \quad E\left\{\sum_{j=1}^{n} c_j(x_j + e_j)\right\}$$

$$\text{subject to} \quad x \in \Omega$$

$$\text{Prob}\left\{\sum_{j=1}^{n} A_{ij}(x_j + e_j) \geq b_i\right\} \geq \gamma_i \qquad i = 1, \ldots, m$$

where A , b , c , γ are given, A is an $m \times n$ matrix, c an $n \times 1$ vector, b and γ are $m \times 1$ vectors with the following restriction on γ : $0 \leq \gamma_i \leq 1$ for $i = 1, \ldots, m$; x is an $n \times 1$ vector of decisions and e is an $n \times 1$ vector of random variables. Ω is a polyhedral set defined by a set of linear constraints.

C. Model II

Although in Model I the uncertainty e_j is independent of the decision x_j , this might not always be a desirable feature. This is why we introduce a second model. Here the uncertainty is proportional to the decision. For an a priori decision x_j we have an outcome $x_j + \alpha_j x_j$ where α_j is a random variable. This can be used to model spoilage in the diet problem where x_j are the amounts of food ordered and $\alpha_j x_j$ the amounts spoiled that cannot be used at consumption time (α_j would have value between -1 and 0 in this example). We have now the following model.

Model II:

$$\text{minimize} \quad E\left\{\sum_{j=1}^{n} c_j(1+\alpha_j)x_j\right\}$$

$$\text{subject to} \quad x \in \Omega$$

$$\text{Prob}\left\{\sum_{j=1}^{n} A_{ij}(1+\alpha_j)x_j \geq b_i\right\} \geq \gamma_i \qquad i = 1, \ldots, m$$

where everything is defined as in Model I except for α which is an $n \times 1$ vector of random variables.

Model II overcomes the shortcomings mentioned for Model I. However, we had to pay the price in increased complexity. This complexity will make Model II much less tractable mathematically than Model I.

We will now present schemes to solve Models I and II.

D. Deterministic Equivalents for Models I and II

We will use here the approach developed by Charnes and Cooper in [3], [4]. In order to solve these chance-constrained problems we will find what is known in the literature as "deterministic equivalents" to these problems. A deterministic equivalent problem is a reformulation of the original chance-constrained problem where all random elements have been eliminated. Deterministic equivalents for Models I and II will be presented in Theorems 1.1 and 1.4. However we will first need the following definitions.

Definition:

The distribution function of a random variable X is the function $G(z)$ defined as $G(z) = \text{Prob}\{X \leq z\}$.

While the definition of the distribution function presents no difficulty, we need to define the inverse of this function for the general case.

Definition:

The inverse of the distribution function $G(z)$ of a random variable X is defined as follows:

$$G^{-1}(\gamma) = \sup \{z \mid G(z) \leq \gamma\}$$

for $0 \leq \gamma \leq 1$ and it is undefined elsewhere.

We will also need the following functions.

Definition:

We will call the "tilde" distribution function of a random variable X the following function:

$$\tilde{G}(z) = \text{Prob}\{X < z\} .$$

Its inverse is defined as:

$$\tilde{G}^{-1}(\gamma) = \sup \{z \mid \tilde{G}(z) \leq \gamma\}$$

for $0 \leq \gamma \leq 1$ and is undefined elsewhere.

Note:

1. If X is a continuous random variable then $G(z) = \text{Prob}\{X \leq z\} = \text{Prob}\{X < z\} = \tilde{G}(z)$ and $G^{-1}(\gamma) = \tilde{G}^{-1}(\gamma)$ for $0 \leq \gamma \leq 1$.

2. If X is a discrete random variable then $\tilde{G}(z) = \text{Prob}\{X < z\} = \text{Prob}\{X \leq z - 1\} = G(z - 1)$.

We will now proceed to define the deterministic equivalents mentioned earlier.

1. Model I

Theorem 1.1:

The following linear program is a deterministic equivalent to Model I:

$$\text{minimize} \quad \sum_{j=1}^{n} c_j(x_j + E(e_j))$$

$$\text{subject to} \quad x \in \Omega$$

$$\sum_{j=1}^{n} A_{ij}x_j \geq b_i - \tilde{G}_i^{-1}(1 - \gamma_i) \qquad i = 1, \ldots, m$$

where $\tilde{G}_i(z) = \text{Prob}\left\{\sum_{j=1}^{n} A_{ij}e_j < z\right\}$.

Proof:

For all $i = 1, \ldots, m$

$$\text{Prob}\left\{\sum_{j=1}^{n} A_{ij}(x_j + e_j) \geq b_i\right\} \geq \gamma_i$$

$$\Longleftrightarrow \text{Prob}\left\{\sum_{j=1}^{n} A_{ij}e_j \geq b_i - \sum_{j=1}^{n} A_{ij}x_j\right\} \geq \gamma_i$$

$$\Longleftrightarrow 1 - \text{Prob}\left\{\sum_{j=1}^{n} A_{ij}e_j < b_i - \sum_{j=1}^{n} A_{ij}x_j\right\} \geq \gamma_i$$

$$\Longleftrightarrow 1 - \tilde{G}_i\left[b_i - \sum_{j=1}^{n} A_{ij}x_j\right] \geq \gamma_i$$

$$\Longleftrightarrow \tilde{G}_i\left[b_i - \sum_{j=1}^{n} A_{ij}x_j\right] \leq 1 - \gamma_i$$

$$\Longleftrightarrow b_i - \sum_{j=1}^{n} A_{ij}x_j \leq \tilde{G}^{-1}(1 - \gamma_i)$$

$$\Longleftrightarrow \sum_{j=1}^{n} A_{ij}x_j \geq b_i - \tilde{G}^{-1}(1 - \gamma_i) \ . \blacksquare$$

<u>Corollary 1.2</u>:

If in Model I the random variables are continuous, then the following linear program is a deterministic equivalent for Model I.

$$\text{minimize} \quad \sum_{j=1}^{n} c_j(x_j + E(e_j))$$

$$\text{subject to} \quad x \in \Omega$$

$$\sum_{j=1}^{n} A_{ij}x_j \geq b_i - G_i^{-1}(1 - \gamma_i) \qquad i = 1, \ldots, m$$

where $G_i(z) = \text{Prob}\left\{\sum_{j=1}^{n} A_{ij}e_j \leq z\right\}$.

<u>Proof</u>:

This is Theorem 1.1 with the observation that $G_i^{-1}(\gamma) = \tilde{G}_i(\gamma)$ for $0 \leq \gamma \leq 1$ when e_j $j = 1, \ldots, n$ are continuous. ■

<u>Corollary 1.3</u>:

If in Model I the random variables are discrete, then the following linear program is a deterministic equivalent for Model I.

$$\text{minimize} \quad \sum_{j=1}^{n} c_j(x_j + E(e_j))$$

$$\text{subject to} \quad x \in \Omega$$

$$\sum_{j=1}^{n} A_{ij}x_j \geq b_i - G_i'^{-1}(1 - \gamma_i) \qquad i = 1, \ldots, m$$

where $G_i'(z) = G_i(z - 1) = \text{Prob}\left\{\sum_{j=1}^{n} A_{ij}e_j \leq z - 1\right\}$.

<u>Proof</u>:

This is Theorem 1.1 with the observation that in this case $\tilde{G}_i(z) = G_i(z - 1)$. ■

2. Model II

<u>Theorem 1.4</u>:

The following nonlinear program is a deterministic equivalent to Model II.

$$\text{minimize} \quad \sum_{j=1}^{n} c_j(1 + E(\alpha_j))x_j$$

$$\text{subject to} \quad x \in \Omega$$

$$\sum_{j=1}^{n} A_{ij}x_j + {}_i\tilde{G}_x^{-1}(1 - \gamma_i) \geq b_i \qquad i = 1, \ldots, m$$

where ${}_i\tilde{G}_x(z) = \text{Prob}\left\{\sum_{j=1}^{n} A_{ij}x_j\alpha_j < z\right\}$.

Proof:

For all $i = 1, \ldots, m$

$$\text{Prob}\left\{\sum_{j=1}^{n} A_{ij}(1 + \alpha_j)x_j \geq b_i\right\} \geq \gamma_i$$

$$\Longleftrightarrow \text{Prob}\left\{\sum_{j=1}^{n} A_{ij}\alpha_j x_j \geq b_i - \sum_{j=1}^{n} A_{ij}x_j\right\} \geq \gamma_i$$

$$\Longleftrightarrow 1 - \text{Prob}\left\{\sum_{j=1}^{n} A_{ij}\alpha_j x_j < b_i - \sum_{j=1}^{n} A_{ij}x_j\right\} \geq \gamma_i$$

$$\Longleftrightarrow 1 - {}_i\tilde{G}_x\left[b_i - \sum_{j=1}^{n} A_{ij}x_j\right] \geq \gamma_i$$

$$\Longleftrightarrow {}_i\tilde{G}_x\left[b_i - \sum_{j=1}^{n} A_{ij}x_j\right] \leq 1 - \gamma_i$$

$$\Longleftrightarrow b_i - \sum_{j=1}^{n} A_{ij}x_j \leq {}_i\tilde{G}_x^{-1}(1 - \gamma_i)$$

$$\Longleftrightarrow \sum_{j=1}^{n} A_{ij}x_j + {}_i\tilde{G}_x^{-1}(1 - \gamma_i) \geq b_i \; . \blacksquare$$

Corollary 1.5:

If in Model II the random variables are continuous then the following nonlinear program is a deterministic equivalent for Model II.

$$\text{minimize} \quad \sum_{j=1}^{n} c_j(1 + E(\alpha_j))x_j$$

$$\text{subject to} \quad x \in \Omega$$

$$\sum_{j=1}^{n} A_{ij}x_j + {}_iG_x^{-1}(1 - \gamma_i) \geq b_i \qquad i = 1, \ldots, m$$

where ${}_iG_x(z) = \text{Prob}\left\{\sum_{j=1}^{n} A_{ij}\alpha_j x_j \leq z\right\}$.

Proof:

By Theorem 1.4 and using the fact that here ${}_i\tilde{G}_x(z) = {}_iG_x(z)$ and ${}_i\tilde{G}_x^{-1}(\gamma) = {}_iG_x^{-1}(\gamma)$ for $0 \leq \gamma \leq 1$. ■

Corollary 1.6:

If in Model II the random variables are discrete, then the following nonlinear program is a deterministic equivalent to Model II.

$$\text{minimize} \quad \sum_{j=1}^{n} c_j(1 + E(\alpha_j))x_j$$

$$\text{subject to} \quad x \in \Omega$$

$$\sum_{j=1}^{n} A_{ij}x_j + {}_iG_x'^{-1}(1 - \gamma_i) \geq b_i \qquad i = 1, \ldots, m$$

where

$${}_iG'_x(z) = {}_iG_x(z - 1)$$

and

$${}_iG_x(z) = \text{Prob}\left\{\sum_{j=1}^{n} A_{ij}\alpha_j x_j \le z\right\} .$$

Proof:

By Theorem 1.4 and using the fact that ${}_i\tilde{G}_x(z) = {}_iG_x(z - 1)$.■

E. Limitations of the Deterministic Equivalents of Models I and II

The deterministic equivalents presented in the preceding section assume that the convolutions

$$\text{(i)} \quad G_i(z) = \text{Prob}\left\{\sum_{j=1}^{n} A_{ij}e_j \le z\right\}$$

$$\text{(ii)} \quad {}_iG_x(z) = \text{Prob}\left\{\sum_{j=1}^{n} A_{ij}\alpha_j x_j \le z\right\}$$

can be calculated and the inverse can be determined exactly for (i), and in terms of x in (ii). Although this can be done easily when the random variables are independent and have normal distributions, as will be shown in Chapter III, in general it is not the case. These convolutions in the case of (i) can be very difficult to calculate or in the case of (ii) it might be difficult to express the inverse distribution explicitly in one simple expression of x. We will also point out that in practice the distributions of the random variables themselves might not be completely known. So far the discussion has been concerned with determining precisely the coefficients of the constraints; however, another difficulty arises: If we succeed in calculating these convolutions and their inverses, Model I is a simple linear program and can be solved. This is not the case for Model II. We have then a nonlinear program which might not only be difficult to formulate but also to solve.

This is why we look for other ways of solving Models I and II. We will concentrate on obtaining what we call conservative approximation problems for these models.

Definition:

A problem B is called a conservative approximation of a problem A if and only if:

(i) The feasibility set of B is contained in the feasibility set of A.

(ii) A and B have the same objective function.

In the next chapter we will present a set C of random variables for which conservative approximations for Models I and II cannot only be easily constructed with limited information about the random variables, but also are linear programming problems.

CHAPTER II

In this chapter we present a set of random variables C and a corresponding class $\mathcal{C}$ of densities and we discuss their properties. These will be useful for defining conservative approximations for Models I and II.

The set C consists of the continuous random variables which have densities that are symmetric, unimodal with maximum at 0, and with finite range. Examples of such distributions are the truncated normal, the uniform distribution, certain Beta distributions, the truncated double-exponential, and many others. In the context of our models such distributions can be used to model many real situations. The fact that the distributions have finite range is certainly realistic. The symmetry and being unimodal with maximum at the mean are more restrictive assumptions; however, anything that can be modeled as a truncated normal would fit. As an example, measurement errors can certainly be modeled using a truncated normal. Spoilage is another example which could be modeled if the quantities spoiled tend to have distributions concentrated around their means. We will now turn to the study of this class of distributions.

A. The Class $\mathcal{C}$ of Densities and the Set of Random Variables C

Definition 2.1:

$\mathcal{C}$ is the class of densities f such that $f \in \mathcal{C}$ if and only if

a) f is symmetric around $0 : \forall x \in (-\infty,+\infty) \quad f(x) = f(-x)$.

b) f is unimodal with maximum at 0.

c) f has finite range: $\exists a$, $0 < a < +\infty$ such that

$f(x) = 0$ for $x \notin [-a,+a]$.

d) f is continuous.

Notation:

The distribution function of a random variable having density f is denoted by $F(x) = \int_{-\infty}^{x} f(u)du$.

This class C has well-known properties that can be derived from the symmetry of the densities.

Proposition 2.2:

If X is a random variable with density $f \in C$ then the following properties hold:

(i) $E(X) = 0$

(ii) $F(x) = 1 - F(-x) \quad \forall x \in (-\infty,+\infty)$

(iii) $F(0) = \frac{1}{2}$.

Proof:

(i) $E(X) = \int_{-\infty}^{+\infty} xf(x)dx = \int_{-\infty}^{0} xf(x)dx + \int_{0}^{+\infty} xf(x)dx$

changing variables in the second integral to $u = -x$

$$\Rightarrow E(X) = \int_{-\infty}^{0} xf(x)dx + \int_{0}^{-\infty} (-u)f(-u)(-du)$$

$$\Rightarrow E(X) = \int_{-\infty}^{0} xf(x)dx - \int_{0}^{-\infty} xf(-x)dx .$$

However by a) of Definition 2.1 $f(x) = f(-x) \quad \forall x \in (-\infty,\infty)$

$$\Rightarrow E(X) = \int_{-\infty}^{0} (x - x)f(x)dx = 0 .$$

(ii) $F(x) = \int_{-\infty}^{x} f(u)du$ changing variables to $v = -u$

$$\Rightarrow F(x) = \int_{+\infty}^{-x} f(-v)(-dv)$$

$$\Rightarrow F(x) = \int_{-x}^{+\infty} f(-v)dv .$$

By symmetry of f : $f(-v) = f(v) \quad \forall v$

$$\Rightarrow F(x) = \int_{-x}^{+\infty} f(v)dv$$

$$= \int_{-\infty}^{+\infty} f(v)dv - \int_{-\infty}^{-x} f(v)dv$$

$$= 1 - F(-x) .$$

(iii) By (ii)

$$F(0) = 1 - F(0)$$

$$\Rightarrow 2F(0) = 1$$

$$\Rightarrow F(0) = \frac{1}{2} . \blacksquare$$

Definition 2.3:

For a density of the class C the inverse of the distribution function $F(x)$ is defined to be:

$$F^{-1}(0) = \max \{z \mid F(z) = 0\}$$

$$F^{-1}(1) = \min \{z \mid F(z) = 1\}$$

and for $0 < \gamma < 1$

$$F^{-1}(\gamma) = \{z \mid F(z) = \gamma\} .$$

Using this definition more properties of C are presented below:

Proposition 2.4:

For any density $f \in C$ having range $[-a,+a]$ the following holds:

(i) $F^{-1}(0) = -a$ and $F^{-1}(1) = +a$

(ii) $F^{-1}\left(\frac{1}{2}\right) = 0$.

Proof:

(i) Apply Definition 2.3.

(ii) Apply (iii) of Proposition 2.2.

Theorem 2.5:

If X is a random variable with density $f \in C$ then for all real numbers c , cX is a random variable with density belonging to C .

Proof:

Call $g(x)$ the density of cX . Then $g(x) = f\left(\frac{x}{c}\right)$.

a) $g(x) = g(-x) \ \forall x \in (-\infty,+\infty)$ since $g(x) = f\left(\frac{x}{c}\right) = f\left(-\frac{x}{c}\right) = f\left(\frac{-x}{c}\right) = g(-x)$.

b) *Case 1*: $c \geq 0$ as $g(x) = f\left(\frac{x}{c}\right)$ $g(x)$ is unimodal with maximum at 0.

Case 2: $c < 0$ $g(x) = f\left(\frac{x}{c}\right) = f\left(\frac{x}{-c}\right)$ by symmetry of f and this is the same as Case 1.

c) $g(x) = f\left(\frac{x}{c}\right)$ has finite range $\left[-\frac{a}{c}, +\frac{a}{c}\right]$ for $c \geq 0$ or $\left[\frac{a}{c}, -\frac{a}{c}\right]$ for $c < 0$.

Definition 2.6:

Call C the set of random variables which have a density belonging to $\mathcal{C}$.

Theorem 2.7:

If $X \in C$ and has a distribution function F and cX has a distribution function G (c real number) then:

$$\forall \gamma \in [0,1]$$

$$G^{-1}(\gamma) = |c| F^{-1}(\gamma) .$$

Proof:

$\forall \gamma \in [0,1] \ \exists z$ such that

$$z = G^{-1}(\gamma)$$

and

$$G(z) = \gamma$$

$$\Rightarrow \gamma = \text{Prob}\ \{cX \le z\}\ .$$

Case 1: $c > 0$

$$\Rightarrow \gamma = \text{Prob}\left\{X \le \frac{z}{c}\right\}$$

$$\Rightarrow \gamma = F\left(\frac{z}{c}\right)$$

$$\Rightarrow F^{-1}(\gamma) = \frac{z}{c}$$

$$\Rightarrow z = cF^{-1}(\gamma)$$

$$\Rightarrow G^{-1}(\gamma) = cF^{-1}(\gamma)\ .$$

Case 2: $c < 0$

$$\gamma = \text{Prob}\left\{X \ge \frac{z}{c}\right\}$$

$$= 1 - \text{Prob}\left\{X \le \frac{z}{c}\right\}$$

$$= 1 - F\left(-\frac{z}{|c|}\right)$$

since $X \varepsilon C$

$$= F\left(\frac{z}{|c|}\right)$$

$$\Rightarrow F^{-1}(\gamma) = \frac{z}{|c|}$$

$$\Rightarrow G^{-1}(\gamma) = |c|F^{-1}(\gamma)\ .\ \blacksquare$$

Theorem 2.8:

If $f_j \varepsilon C$, $j = 1,2$, then the convolution $f_1 * f_2 \varepsilon C$.

Proof: $\forall f_1$ and $f_2 \varepsilon C$

$$g = f_1 * f_2 .$$

a) g is symmetric

$$\forall z \ \varepsilon \ (-\infty,\infty)$$

$$g(z) = \int_{-\infty}^{+\infty} f_1(z - u)f_2(u)du$$

changing variables $u = -v$

$$= \int_{+\infty}^{-\infty} f_1(z + v)f_2(-v)(-dv)$$

$$= \int_{-\infty}^{+\infty} f_1(z + v)f_2(-v)dv$$

since f_1 and f_2 are symmetric

$$g(z) = \int_{-\infty}^{+\infty} f_1(-z - v)f_2(v)dv$$

$$= g(-z) .$$

b) It is sufficient to show that $g(z)$ is increasing in z for $z \leq 0$ or $\forall z \leq 0$ and $\delta z > 0$ such that $z + \delta z \leq 0$

$$g(z + \delta z) - g(z) \geq 0 .$$

Then by symmetry $g(z)$ is unimodal with maximum at 0.

$$g(z + \delta z) - g(z)$$

$$= \int_{-\infty}^{+\infty} f_1(z + \delta z - u) f_2(u) du - \int_{-\infty}^{+\infty} f_1(z - u) f_2(u) du$$

$$= \int_{-\infty}^{z + \frac{1}{2}\delta z} [f_1(z + \delta z - u) - f_1(z - u)] f_2(u) du$$

$$+ \int_{z + \frac{1}{2}\delta z}^{+\infty} [f_1(z + \delta z - u) - f_1(z - u)] f_2(u) du .$$

Changing variable in the first integral to $v = 2z + \delta z - u$

$$g(z + \delta z) - g(z)$$

$$= - \int_{+\infty}^{z + \frac{1}{2}\delta z} [f_1(-z + u) - f_1(-z - \delta z + u)] f_2(2z + \delta z - u) du$$

$$+ \int_{z + \frac{1}{2}\delta z}^{+\infty} [f_1(z + \delta z - u) - f_1(z - u)] f_2(u) du .$$

Since f_1 is symmetric

$$f_1(-z + u) = f_1(z - u)$$

$$f_1(-z - \delta z + u) = f_1(z + \delta z - u)$$

$$\Rightarrow g(z + \delta z) - g(z)$$

$$= \int_{z+\frac{1}{2}\delta z}^{+\infty} [f_1(z + \delta z - u) - f_1(z - u)](f_2(u) - f_2(2z + \delta z - u))du .$$

The proof will now proceed as follows. We shall show that:

1) $f_1(z + \delta z - u) - f_1(z - u) \geq 0$ for $z + \frac{1}{2}\delta z \leq u \leq +\infty$

2) $f_2(u) - f_2(2z + dz - u) \geq 0$ for $z + \frac{1}{2}\delta z \leq u \leq +\infty$.

Then it is clear that the whole integral is nonnegative and $g(z + \delta z) - g(z) \geq 0$.

1) *Case 1:*

$$z + \frac{1}{2}\delta z \leq u \leq z + \delta z$$

$$\Rightarrow -z - \delta z \leq -u \leq -z - \frac{1}{2}\delta z$$

$$\Rightarrow \left\{ \begin{array}{c} -\delta z \leq z - u \leq -\frac{1}{2}\delta z \\ \text{and} \\ 0 \leq z + \delta z - u \leq \frac{1}{2}\delta z \end{array} \right\}$$

$$\Rightarrow 0 \leq z + \delta z - u \leq \frac{1}{2}\delta z \leq -(z - u) .$$

As $f_1 \varepsilon C$, f_1 is symmetric and unimodal with maximum at 0.

$$f_1(z + \delta z - u) \geq f_1(-(z - u)) = f_1(z - u)$$

$$\Rightarrow f_1(z + \delta z - u) - f_1(z - u) \geq 0 .$$

Case 2:

$$z + \delta z \leq u \leq +\infty .$$

Then

$$z + \delta z - u \leq 0$$

as $\delta z > 0$

$$z - u \leq z + \delta z - u \leq 0$$

since $f_1 \varepsilon C$ $f_1(x)$ is increasing for $x \leq 0$ and $f_1(z + \delta z - u) - f_1(z - u) \geq 0$.

2) *Case 1:*

$$z + \frac{1}{2}\delta z \leq u \leq 0$$

$$\Rightarrow 0 \leq -u \leq -z - \frac{1}{2}\delta z$$

$$\Rightarrow 2z + \delta z \leq 2z + \delta z - u \leq z + \frac{1}{2}\delta z$$

$$\Rightarrow 2z + \delta z - u \leq z + \frac{1}{2}\delta z \leq u \leq 0$$

as $f_2 \varepsilon C$, $f_2(x)$ is increasing for $x \leq 0$

$$\Rightarrow f_2(u) - f_2(2z + \delta z - u) \geq 0 .$$

Case 2:

$$u \leq 0 \leq +\infty$$

$$f_2(u) - f_2(2z + \delta z - u) = f_2(u) - f_2(u - 2z - \delta z)$$

since f_2 is symmetric.

However

$$2z + \delta z \leq 0$$

$$\Rightarrow 0 \leq u \leq u - 2z - \delta z$$

as $f_2 \varepsilon C$ $f_2(x)$ is decreasing for $x \geq 0$

$$\Rightarrow f_2(u) - f_2(u - 2z - \delta z) \geq 0 .$$

c) If f_j , $j = 1,2$, have ranges $[-a_j,+a_j]$ then $f_1 * f_2$ has a finite range:

$$[-a_1 - a_2, a_1 + a_2]$$

as $(f_1 * f_2)(z) = 0$ for $z \notin [-a_1 - a_2, a_1 + a_2]$ since $f_j(z) = 0$ for $z \notin [-a_j, a_j]$ for $j = 1,2$.

d) Since f_1 and f_2 are continuous,

$$f_1 * f_2 = \int_{-\infty}^{+\infty} f_1(z - u) f_2(u) du$$

is continuous.

<u>Note</u>:

An alternate proof for b) is outlined in a note on page 164 of William Feller's book [7].■

Corollary 2.9:

If X_j $j = 1, \ldots, n$ are independent and such that $X_j \varepsilon C$ $j = 1, \ldots, n$ then $\sum_{j=1}^{n} X_j \varepsilon C$.

Proof:

- True for $n = 1$
- True for $n = 2$ by Theorem 2.8
- Assume it is true for $n = k$

i.e., $g_k = \mathop{*}_{j=1}^{k} f_j$, f_j density of X_j .

Then $g_k \varepsilon C$.

However $g_k * f_{k+1} = g_{k+1} \varepsilon C$ by Theorem 2.8. Therefore it is true for $k + 1$ and true for all n .

Corollary 2.10:

If the independent random variables X_j $j = 1, \ldots, n$ belong to C , then for all real numbers c_j $j = 1, \ldots, n$: $\sum_{j=1}^{n} c_j X_j \varepsilon C$.

Proof:

Apply Theorem 2.5 and Corollary 2.9.■

B. Bounds on the Distribution Function of the Sum of Random Variables Belonging to C

Theorem 2.11:

If $X_j \varepsilon C$ $j = 1, \ldots, n$ are independent with distribution functions F_j and calling $G = \overset{n}{\underset{j=1}{*}} F_j$ then:

For $-\sum_{j=1}^{n} F_j^{-1}(0) \geq x \geq 0$

$$\frac{x - \sum_{j=1}^{n} F_j^{-1}(0)}{-2 \sum_{j=1}^{n} F_j^{-1}(0)} \leq G(x) .$$

For $+\sum_{j=1}^{n} F_j^{-1}(0) \leq x \leq 0$

$$G(x) \leq \frac{x - \sum_{j=1}^{n} F_j^{-1}(0)}{-2 \sum_{j=1}^{n} F_j^{-1}(0)} .$$

Proof:

By Corollary 2.9 $\sum_{j=1}^{n} X_j \varepsilon C$; therefore, $g = \overset{n}{\underset{j=1}{*}} f_j$ is unimodal with maximum at 0. Therefore $g(x)$ is increasing for $x \leq 0$ and hence $G(x) = \int_{-\infty}^{x} g(u)du$ is convex for $x \leq 0$. Furthermore, we know that

$$G\left(\sum_{j=1}^{n} F_j^{-1}(0)\right) = 0$$

and

$$\sum_{j=1}^{n} F_j^{-1}(0) > -\infty$$

$$G(0) = \frac{1}{2} .$$

Therefore $G(x)$ has for upper bound the straight line from $\left(\sum_{j=1}^{n} F_j^{-1}(0), 0\right)$ to $\left(0, \frac{1}{2}\right)$.

For $x \leq 0$,

$$G(x) \leq \frac{x - \sum_{j=1}^{n} F_j^{-1}(0)}{-2 \sum_{j=1}^{n} F_j^{-1}(0)} .$$

Since $G(x) \varepsilon C$ it is symmetric and by symmetry for $x \geq 0$

$$\frac{x - \sum_{j=1}^{n} F_j^{-1}(0)}{-2 \sum_{j=1}^{n} F_j^{-1}(0)} \leq G(x) . \blacksquare$$

<u>Theorem 2.12</u>:

If the independent random variables X_j $j = 1, \ldots, n$ belong to C , then:

For $x \geq 0$

$$G(x) \leq \min_{j=1,\ldots,n} \{F_j(x)\} .$$

For $x \leq 0$

$$\max_{j=1,\ldots,n} \{F_j(x)\} \leq G(x) .$$

<u>Proof</u>:

It is sufficient to prove that for $x \leq 0$ $F_j(x) \leq G(x)$ $\forall j = 1, \ldots, n$.

<u>Proof by Induction</u>:

- For $n = 1$, trivial $F_j(x) \leq F_j(x)$
- For $n = 2$

For $x \leq 0$

$$G_2(x) - F_j(x) \quad \left(\begin{matrix} j = 1,2 \\ k \neq j \end{matrix} \;,\; G_2(x) = F_1 * F_2(x)\right)$$

$$= \int_{-\infty}^{+\infty} F_j(x - u) f_k(u) du - F_j(x)$$

$$= \int_{-\infty}^{+\infty} F_j(x - u) f_k(u) du - \int_{-\infty}^{+\infty} F_j(x) f_k(u) du$$

since $\int_{-\infty}^{+\infty} f_k(u) du = 1$

$$= \int_{-\infty}^{+\infty} [F_j(x - u) - F_j(x)]f_k(u)du$$

$$= \int_{0}^{+\infty} [F_j(x - u) - F_j(x)]f_k(u)du + \int_{-\infty}^{0} [F_j(x - u) - F_j(x)]f_k(u)du .$$

Changing variables in the second integral $u = -t$

$$= \int_{0}^{+\infty} [F_j(x - u) - F_j(x)]f_k(u)du + \int_{+\infty}^{0} [f_j(x + t) - F_j(x)]f_k(-t)(-dt)$$

since $f_k(u)$ is symmetric:

$$= \int_{0}^{+\infty} [F_j(x - u) + F_j(x + u) - 2F_j(x)]f_k(u)du$$

$f_k(u) \geq 0 \ \forall u$ it is sufficient to show that $[F_j(x - u) + F_j(x + u) - 2F_j(x)] \geq 0$ for $u \in [0,+\infty]$ to insure that $G_2(x) - F_j(x) \geq 0$ and $G_2(x) \geq F_j(x) \quad j = 1,2$.

1) For $u = 0$

$$[F_j(x - 0) + F_j(x + 0) - 2F_j(x)] = 0 .$$

2) For $0 \leq u \leq +\infty$

$$[F_j(x - u) + F_j(x + u) - 2F_j(x)]$$

is an increasing function of u since its derivative is:

$$-f_j(x-u) + f_j(x+u) = f_j(x+u) - f_j(u-x) \geq 0$$

since f_j is symmetric and unimodal and $|u - x| \geq |x + u|$ for $0 \leq u \leq +\infty$.

a) $0 \leq u \leq -x \Rightarrow x \leq x + u \leq 0 \leq -x \leq u - x$

b) $-x \leq u \leq +\infty \Rightarrow 0 \leq x + u \leq u - x$.

Therefore it is true for $n = 2$.

• Assume it is true for $n = k$.

Then for $x \leq 0$

$$F_j(x) \leq G_k(x) \qquad j = 1, \ldots, k .$$

Since C is closed under convolution the density $g_k(x) \varepsilon C$.
Then by the proof for $n = 2$:

For $x \leq 0$

$$G_k(x) \leq G_k * F_{k+1}(x) = G_{k+1}(x)$$

$$F_{k+1}(x) \leq G_k * F_{k+1}(x) = G_{k+1}(x) .$$

Therefore:

For $x \leq 0$

$$F_j(x) \leq G_k(x) \qquad \text{for } j = 1, \ldots, k+1 .$$

It is true for all n . ■

Theorem 2.13:

For all $X_j \in C$ $j = 1, \ldots, n$ independent random variables with distribution function F_j and all real numbers c_j $j = 1, \ldots, n$, $G(x) = \text{Prob}\left\{\sum_{j=1}^{n} c_j X_j \leq x\right\}$ is such that for $-\sum_{j=1}^{n} |c_j| F_j^{-1}(0) \geq x \geq 0$

$$\frac{x - \sum_{j=1}^{n} |c_j| F_j^{-1}(0)}{-2 \sum_{j=1}^{n} |c_j| F_j^{-1}(0)} \leq G(x) \leq \min_{j=1,\ldots,n} \{|c_j| F_j(x)\}$$

and for $\sum_{j=1}^{n} |c_j| F_j^{-1}(0) \leq x \leq 0$

$$\max_{j=1,\ldots,n} \{|c_j| F_j(x)\} \leq G(x) \leq \frac{x - \sum_{j=1}^{n} |c_j| F_j^{-1}(0)}{-2 \sum_{j=1}^{n} |c_j| F_j^{-1}(0)} .$$

Proof:

Apply Theorem 2.12 and 2.11 and Corollary 2.10 and Theorem 2.7.■

Corollary 2.14:

For all $X_j \in C$ $j = 1, \ldots, n$ independent random variables with ranges $[-a_j, +a_j]$ and all real numbers c_j $j = 1, \ldots, n$, $G(x) = \text{Prob}\left\{\sum_{j=1}^{n} c_j X_j \leq x\right\}$ is such that:

For $\sum_{j=1}^{n} |c_j| a_j \geq x \geq 0$

$$\frac{x + \sum_{j=1}^{n} |c_j| a_j}{2 \sum_{j=1}^{n} |c_j| a_j} \le G(x)$$

and for $-\sum_{j=1}^{n} |c_j| a_j \le x \le 0$

$$G(x) \le \frac{x + \sum_{j=1}^{n} |c_j| a_j}{2 \sum_{j=1}^{n} |c_j| a_j} .$$

Proof:

Apply Theorem 2.13 with $F_j^{-1}(0) = -a_j$ from Proposition 2.4. ■

Theorem 2.15:

For all $X_j \in C$ $j = 1, \ldots, n$ independent with ranges $[-a_j, +a_j]$ and real numbers c_j $j = 1, \ldots, n$, if $G(x) = \text{Prob}\left\{\sum_{j=1}^{n} c_j X_j \le x\right\}$ it is true that:

For $0 \le \gamma \le \frac{1}{2}$

$$(2\gamma - 1) \sum_{j=1}^{n} |c_j| a_j \le G^{-1}(\gamma) \le \min_{j=1,\ldots,n} \left\{|c_j| F_j^{-1}(\gamma)\right\} .$$

For $1 \ge \gamma \ge \frac{1}{2}$

$$\max_{j=1,\ldots,n} \left\{|c_j| F_j^{-1}(\gamma)\right\} \le G^{-1}(\gamma) \le (2\gamma - 1) \sum_{j=1}^{n} |c_j| a_j .$$

Proof:

This is just a different way of stating Theorem 2.13. As $X_j \varepsilon C$, all the inverses of F_j are concave and so is $G^{-1}(\gamma)$ and $(2\gamma - 1) \sum_{j=1}^{n} |c_j| a_j$ is just the straight line joining $G^{-1}(0)$, $G^{-1}(1)$ and going through $G^{-1}\left(\frac{1}{2}\right)$. The concavity of $G^{-1}(\gamma)$ explains the inequalities.

As both F_j and G are strictly increasing in their ranges as $X_j \varepsilon C$ for all $j = 1, \ldots, n$, then it is clear that

$$\forall x \varepsilon (-\infty, 0] \quad |c_j| F_j^{-1}(x) \leq G(x) .$$

Then $\gamma \varepsilon \left[0, \frac{1}{2}\right]$, $|c_j| F_j^{-1}(\gamma) \geq G^{-1}(\gamma)$. ■

CHAPTER III

In this chapter the results of the previous chapter are used to find conservative approximations to Models I and II. The normal case is treated separately at the end of the chapter.

A. A Conservative Approximation to Model I

Theorem 3.1:

If $e_j \in C$ $j = 1, \ldots, n$ with ranges $[-a_j, +a_j]$, the following linear program is a conservative approximation to Model I for $\gamma \geq \frac{1}{2}$.

$$\text{minimize} \quad c^T x$$

$$\text{subject to} \quad x \in \Omega$$

$$\sum_{j=1}^{n} A_{ij} x_j \geq b_i + (2\gamma_i - 1) \sum_{j=1}^{n} |A_{ij}| a_j \qquad i = 1, \ldots, m .$$

Proof:

(i) The objective function is $c^T x$ since $\forall j$ $E(e_j) = 0$ as $e_j \in C$.

(ii) From Theorem 2.15 it is clear that for $G_i(z) = \text{Prob}\left\{\sum_{j=1}^{n} A_{ij} e_j \leq z\right\}$ and $\gamma_i \geq \frac{1}{2}$ $\forall i = 1, \ldots, m$

$$(2(1 - \gamma_i) - 1) \sum_{j=1}^{n} a_j |A_{ij}| \leq G_i^{-1}(1 - \gamma_i)$$

$$\Longleftrightarrow (1 - 2\gamma_i) \sum_{j=1}^{n} a_j |A_{ij}| \leq G_i^{-1}(1 - \gamma_i) .$$

Using the deterministic equivalent of Theorem 1.1 it is true that if x is such that for all $i = 1, \ldots, m$,

$$\sum_{j=1}^{n} A_{ij}x_j \geq b_i + (2\gamma_i - 1) \sum_{j=1}^{n} a_j|A_{ij}| .$$

Then

$$\sum_{j=1}^{n} A_{ij}x_j \geq b_i - G_i^{-1}(1 - \gamma_i)$$

since $(2\gamma_i - 1) \sum_{j=1}^{n} a_j|A_{ij}| \geq -G_i^{-1}(1 - \gamma_i)$ and therefore:

$$\text{Prob}\left\{\sum_{j=1}^{n} A_{ij}(x_j + e_j) \geq b_i\right\} \geq \gamma_i \quad \text{for all } i .$$

Hence it is a conservative approximation to Model I. ∎

B. A Conservative Approximation to Model II

Theorem 3.2:

If the Ω of Model II contains the set $\{x \mid x \geq 0\}$ and if $\alpha_j \in C$ $j = 1, \ldots, n$ with ranges $[-a_j, +a_j]$, the following linear program is a conservative approximation to Model II for $\gamma \geq \frac{1}{2}$.

$$\text{minimize} \quad c^T x$$

$$\text{subject to} \quad x \in \Omega$$

$$\sum_{j=1}^{n} [A_{ij} + (1 - 2\gamma_i)|A_{ij}|a_j]x_j \geq b_i \qquad i = 1, \ldots, m .$$

<u>Proof</u>:

(i) The objective function is $c^T x$ since $E\{\alpha_j\} = 0$ for $j = 1, \ldots, n$ as $\alpha_j \in C$.

(ii) If we call

$$_iG_x(z) = \text{Prob}\left\{\sum_{j=1}^{n} A_{ij}x_je_j \leq z\right\}$$

Using again Theorem 2.15 and the fact that $x \geq 0$.

$$\text{for } \gamma_i \geq \frac{1}{2} \quad i = 1, \ldots, n$$

$$(2(1-\gamma_i) - 1) \sum_{j=1}^{n} a_j|A_{ij}|x_j \leq {}_iG_x^{-1}(1-\gamma_i)$$

$$\iff (1 - 2\gamma_i) \sum_{j=1}^{n} a_j|A_{ij}|x_j \leq {}_iG_x^{-1}(1-\gamma_i) .$$

Using the deterministic equivalent of Theorem 1.2 it is true that if x is such that for all $i = 1, \ldots, m$:

$$\sum_{j=1}^{n} [A_{ij} + (1-2\gamma_i)|A_{ij}|a_j]x_j \geq b_i .$$

Then

$$\sum_{j=1}^{n} A_{ij}x_j + (1-2\gamma_i) \sum_{j=1}^{n} |A_{ij}|a_jx_j \geq b_i$$

and

$$\sum_{j=1}^{n} A_{ij}x_j + {}_iG_x^{-1}(1 - 2\gamma_i)$$

$$\geq \sum_{j=1}^{n} A_{ij}x_j + (1 - 2\gamma_i) \sum_{j=1}^{n} |A_{ij}| a_j x_j \geq b_i$$

and finally

$$\text{Prob} \left\{ \sum_{j=1}^{n} A_{ij}(1 + \alpha_j)x_j \geq b_i \right\} \geq \gamma_i \qquad \text{for all} \quad i \ .$$

Thus, this is a conservative approximation to Model II. ■

C. Special Cases for $\gamma = 1$ and $\gamma = \frac{1}{2}$

These two models presented in Theorems 3.1 and 3.2 are linear programs and thus can be solved using the simplex method. They do not yield the exact solution, but in the absence of a method to obtain the exact solution these are valuable problems as they give us a *feasible solution set and an optimum for that solution set.* It is also to be noted that the solutions are exact for $\gamma = 1$ and $\gamma = \frac{1}{2}$.

For $\gamma = 1$, since the random variables have finite range, the problem is reduced to a linear program with the random variable having their "worst" values. This yields the programs:

Model I:

$$\text{minimize} \quad c^T x$$

$$\text{subject to} \quad x \in \Omega$$

$$\sum_{j=1}^{n} A_{ij}x_j \geq b_i + \sum_{j=1}^{n} |A_{ij}| a_j \qquad i = 1, \ldots, m \ .$$

Model II:

$$\text{minimize} \quad c^T x$$
$$\text{subject to} \quad x \in \Omega$$
$$\sum_{j=1}^{n} (A_{ij} - |A_{ij}| a_j) x_j \geq b_i \qquad i = 1, \ldots, m$$

when $\{x \mid x \geq 0\} \subset \Omega$.

For $\gamma = \frac{1}{2}$, since the random variables are symmetric, the problem is reduced to the linear program where all the random variables are set equal to 0. This yields the two identical programs:

Model I:

$$\text{minimize} \quad c^T x$$
$$\text{subject to} \quad x \in \Omega$$
$$Ax \geq b \; .$$

Model II:

$$\text{minimize} \quad c^T x$$
$$\text{subject to} \quad x \in \Omega$$
$$Ax \geq b \; .$$

We will now prove these statements in the following theorem.

Theorem 3.3:

(i) Model I and the problem of Theorem 3.1 are identical for $\gamma = \frac{1}{2}$ and $\gamma = 1$.

(ii) Model II and the problem of Theorem 3.2 are identical for $\gamma = \frac{1}{2}$ and $\gamma = 1$.

Proof:

(i) For $\gamma = 1$

$$(2\gamma_i - 1) \sum_{j=1}^{n} |A_{ij}| a_j = \sum_{j=1}^{n} |A_{ij}| a_j = G_i^{-1}(1) \quad \forall i = 1, \ldots, m$$

and for $\gamma = \frac{1}{2}$

$$(2\gamma_i - 1) \sum_{j=1}^{n} |A_{ij}| a_j = 0 = G_i^{-1}\left(\frac{1}{2}\right) \quad \forall i = 1, \ldots, m .$$

(ii) Similarly

$$(2\gamma_i - 1) \sum_{j=1}^{n} |A_{ij}| a_j x_j = \sum_{j=1}^{n} |A_{ij}| a_j x_j = {}_iG_x^{-1}(1)$$

$$\text{for } \gamma_i = 1 \text{ , } \forall i = 1, \ldots, m \text{ ;}$$

$$(2\gamma_i - 1) \sum_{j=1}^{n} |A_{ij}| a_j x_j = 0 = {}_iG_x^{-1}\left(\frac{1}{2}\right)$$

$$\text{for } \gamma_i = \frac{1}{2} \text{ , } \forall i = 1, \ldots, m \text{ .}$$

Therefore, the constraints are identical.■

Note:

It should be noted that all previous results can be extended to random variable e_j and α_j with nonzero means but for which the random variables $e_j - E(e_j)$ and $\alpha_j - E(\alpha_j)$ belong to C .

D. The Normal Case

The normal distribution does not belong to the class C , however as it is important, it shall be treated separately in this section. The special properties of the normal distribution enables us to formulate workable deterministic equivalents for Models I and II.

Theorem 3.4:

If e_j $j = 1, \ldots, n$ are distributed normally with means μ_j and variance σ_j^2 , then the following linear program is a deterministic equivalent of Model I.

$$\text{minimize} \quad \sum_{j=1}^{n} c_j x_j + \sum_{j=1}^{n} c_j \mu_j$$

$$\text{subject to} \quad x \in \Omega$$

$$\sum_{j=1}^{n} A_{ij} x_j \geq b_i - \sum_{j=1}^{n} A_{ij}\mu_j - \phi^{-1}(1 - \gamma_i)\left(\sum_{j=1}^{n} A_{ij}^2 \sigma_j^2\right)^{\frac{1}{2}}$$

$$i = 1, \ldots, m$$

where $\phi(z)$ is the distribution function of the standardized normal with mean 0 and variance 1.

Proof:

$$\text{Prob}\left\{\sum_{j=1}^{n} A_{ij}(x_j + e_j) \geq b_i\right\} \geq \gamma_i$$

$$\Longleftrightarrow \sum_{j=1}^{n} A_{ij} x_j \geq b_i - G_i^{-1}(1 - \gamma_i)$$

where

$$G_i(z) = \text{Prob}\left\{\sum_{j=1}^{n} A_{ij}e_j \leq z\right\}$$

by Theorem 1.1.

However since e_j have $N(\mu_j,\sigma_j)$ as distributions, then G_i is

$$N\left(\sum_{j=1}^{n} A_{ij}\mu_j \; , \left(\sum_{j=1}^{n} A_{ij}^2\sigma_j^2\right)^{\frac{1}{2}}\right)$$

$$\Longleftrightarrow G_i(z) = \phi\left[\frac{z - \sum_{j=1}^{n} A_{ij}\mu_j}{\left(\sum_{j=1}^{n} A_{ij}^2\sigma_j^2\right)^{\frac{1}{2}}}\right] .$$

Hence if $(1 - \gamma_i) = G_i(z)$

$$1 - \gamma_i = \phi\left[\frac{z - \sum_{j=1}^{n} A_{ij}\mu_j}{\left(\sum_{j=1}^{n} A_{ij}^2\sigma_j^2\right)^{\frac{1}{2}}}\right]$$

and

$$\frac{z - \sum_{j=1}^{n} A_{ij}\mu_j}{\left(\sum_{j=1}^{n} A_{ij}^2\sigma_j^2\right)^{\frac{1}{2}}} = \phi^{-1}[1 - \gamma_i]$$

$$\Longleftrightarrow G_i^{-1}[1 - \gamma_i] = z = \sum_{j=1}^{n} A_{ij}\mu_j + \phi^{-1}[1 - \gamma_i]\left(\sum_{j=1}^{n} A_{ij}^2\sigma_j^2\right)^{\frac{1}{2}} .$$

Hence for all $i = 1, \ldots, m$

$$\text{Prob}\left\{\sum_{j=1}^{n} A_{ij}(x_j + e_j) \geq b_i\right\} \geq \gamma_i$$

$$\Longleftrightarrow \sum_{j=1}^{n} A_{ij}x_j \geq b_i - G_i^{-1}(1 - \gamma_i) \quad \text{by Theorem 1.1}$$

$$\Longleftrightarrow \sum_{j=1}^{n} A_{ij}x_j \geq b_i - \sum_{j=1}^{n} A_{ij}\mu_j - \phi^{-1}[1 - \gamma_i]\left(\sum_{j=1}^{n} A_{ij}^2\sigma_j^2\right)^{\frac{1}{2}} . \blacksquare$$

<u>Theorem 3.5</u>:

If α_j $j = 1, \ldots, n$ are distributed normally with means μ_j and variance σ_j^2, then the following convex program is a deterministic equivalent of Model II.

$$\text{minimize} \quad \sum_{j=1}^{n} c_j(1 + \mu_j)x_j$$

$$\text{subject to} \quad x \in \Omega$$

$$\sum_{j=1}^{n} A_{ij}(1 + \mu_j)x_j + \phi^{-1}[1 - \gamma_i]\left(\sum_{j=1}^{n} A_{ij}^2\sigma_j^2x_j^2\right)^{\frac{1}{2}} \geq b_i .$$

<u>Proof</u>:

$$\text{Prob}\left\{\sum_{j=1}^{n} A_{ij}(1 + \alpha_j)x_j \geq b_i\right\} \geq \gamma_i$$

$$\Longleftrightarrow \sum_{j=1}^{n} A_{ij}x + {}_iG_x^{-1}[1 - \gamma_i] \geq b_i$$

where

$$ {}_iG_x(z) = \text{Prob}\left\{\sum_{j=1}^{n} A_{ij}(1+\alpha_j)x_j \le b_i\right\} $$

by Theorem 1.2.

However since α_j have for distributions $N(\mu_j,\sigma_j)$

$$ {}_iG_x(z) \text{ is normal } N\left(\sum_{j=1}^{n} A_{ij}x_j\mu_j \,,\left(\sum_{j=1}^{n} A_{ij}^2\sigma_j^2x_j^2\right)^{\frac{1}{2}}\right) $$

$$ \Longleftrightarrow {}_iG_x(z) = \phi\left[\frac{z - \sum_{j=1}^{n} A_{ij}x_j\mu_j}{\left(\sum_{j=1}^{n} A_{ij}^2\sigma_j^2x_j^2\right)^{\frac{1}{2}}}\right] . $$

Using the same reasoning as in the proof of Theorem 3.4

$$ {}_iG_x^{-1}[1-\gamma_i] = \sum_{j=1}^{n} A_{ij}\mu_jx_j + \phi^{-1}[1-\gamma_i]\left(\sum_{j=1}^{n} A_{ij}^2\sigma_j^2x_j^2\right)^{\frac{1}{2}} . $$

Hence for all $i = 1, \ldots, m$

$$ \text{Prob}\left\{\sum_{j=1}^{n} A_{ij}(1+\alpha_j)x_j \ge b_i\right\} \ge \gamma_i $$

$$ \Longleftrightarrow \sum_{j=1}^{n} A_{ij}x_j + {}_iG_x^{-1}[1-\gamma_i] \ge b_i $$

$$ \Longleftrightarrow \sum_{j=1}^{n} A_{ij}(1+\mu_j)x_j + \phi^{-1}[1-\gamma_i]\left(\sum_{j=1}^{n} A_{ij}^2\sigma_j^2x_j^2\right)^{\frac{1}{2}} \ge b_i . $$

The program thus obtained is convex since the above expression is concave. This was shown in the literature in [8]. ■

If one does not wish to solve the preceding convex program using the usual convex programming algorithm, a conservative approximation for Model II can be derived in the normal case. This applies if we have the constraint $\{x \mid x \geq 0\}$ included in the set Ω .

Theorem 3.6:

If in Model II, $\{x \mid x \geq 0\} \subset \Omega$ and α_j $j = 1, \ldots, n$ are normally distributed with means μ_j and variances σ_j^2 , then for $\gamma \geq \frac{1}{2}$ the following linear program is a conservative approximation for Model II.

$$\text{minimize} \quad \sum_{j=1}^{n} c_j(1 + \mu_j)x_j$$

$$\text{subject to} \quad x \in \Omega$$

$$\sum_{j=1}^{n} \left[(1 + \mu_j)A_{ij} + \phi^{-1}(1 - \gamma_i)\sigma_j |A_{ij}|\right] x_j \geq b_i \quad \forall i = 1, \ldots, m$$

$$x \geq 0 .$$

Proof:

This can be established using the following two facts:

1) For $x \geq 0$ $\forall i = 1, \ldots, m$

$$\left(\sum_{j=1}^{n} |A_{ij}| x_j \sigma_j\right)^2 \geq \sum_{j=1}^{n} A_{ij}^2 x_j^2 \sigma_j^2$$

$$\Rightarrow \sum_{j=1}^{n} |A_{ij}| \sigma_j x_j \geq \left(\sum_{j=1}^{n} A_{ij}^2 \sigma_j^2 x_j^2\right)^{\frac{1}{2}} .$$

2) For $\gamma \geq \frac{1}{2}$

$$\phi^{-1}[1 - \gamma_i] \leq 0 \quad \forall i \quad i = 1, \ldots, m$$

Hence by 1) and 2)

$$\phi^{-1}[1 - \gamma_i] \sum_{j=1}^{n} |A_{ij}|\sigma_j x_j \leq \phi^{-1}[1 - \gamma_i]\left(\sum_{j=1}^{n} A_{ij}^2 \sigma_j^2 x_j^2\right)^{\frac{1}{2}} .$$

Therefore for $\gamma_i \geq \frac{1}{2}$, for $i = 1, \ldots, m$ if $x \geq 0$ satisfies

$$\sum_{j=1}^{n} \left[(1 + \mu_j)A_{ij} + \phi^{-1}[1 - \gamma_i]|A_{ij}|\sigma_j\right]x_j \geq b_i$$

$$\iff \sum_{j=1}^{n} (1 + \mu_j)A_{ij}x_j + \phi^{-1}[1 - \gamma_i] \sum_{j=1}^{n} |A_{ij}|\sigma_j \geq b_i .$$

Then

$$\sum_{j=1}^{n} (1 + \mu_j)A_{ij}x_j + \phi^{-1}[1 - \gamma_i]\left(\sum_{j=1}^{n} A_{ij}^2 \sigma_j^2 x_j^2\right)^{\frac{1}{2}}$$

$$\geq \sum_{j=1}^{n} (1 + \mu_j)A_{ij}x_j + \phi^{-1}[1 - \gamma_i] \sum_{j=1}^{n} |A_{ij}|\sigma_j x_j \geq b_i$$

$$\Rightarrow \text{Prob}\left\{\sum_{j=1}^{n} A_{ij}(1 + \alpha_j)x_j \geq b_i\right\} \geq \gamma_i .$$

This problem is a conservative approximation to Model II. ∎

CHAPTER IV

A. Sensitivity Analysis

In this chapter we will study the sensitivity of the optimal solutions of Models I and II with respect to the following four characteristics of the model.

1) The cost vector c .
2) The right-hand side b of the constraints.
3) The probability vector γ .
4) The distributions of the random variables (e_j $j = 1, \ldots, m$ for Model I and α_j $j = 1, \ldots, n$ for Model II).

B. The Objective Function and the Right-Hand Side

1. Model I:

Model I presented in either its deterministic equivalent form or conservative approximation form is a linear program and sensitivity analysis on the cost vector or the right-hand side can be performed in the usual manner according to linear programming theory.

When the approximation is used, although we have no guarantee that a change in the approximate problem will reflect exactly a change in the original problem, we can confidently say that in some cases it will give an indication of the effect of that change. For example, if the right-hand side b is modified as to expand the original feasibility set, it is easy to see that the feasibility set of the conservative approximation will also be expanded. Similarly, a change in the cost vector c in the conservative approximation will yield the conservative approximation to the modified original problem. In short, if in the absence of other valid method, we are ready to settle for this conservative approximation solution, then the sensitivity analysis of the conservative approximation will give us results of the same validity as the solution for which we settled.

2. Model II:

Model II in its deterministic equivalent form is a nonlinear program and the sensitivity analysis will have to be performed according to the specific nonlinear program that Model II represents. However, we have noted that we might not be able to obtain this formulation explicitly, in which case we cannot solve the original problem. Here the conservative approximation derived earlier will prove to be useful. In the case the α_j $j = 1, \ldots, n$ belong to the set C and $\gamma \geq \frac{1}{2}$, the conservative approximation was shown to be a linear program in Theorem 3.2 and the usual sensitivity analysis on the cost vector and the right-hand side can be performed using the theory of linear programming. It is important to note that this procedure will yield the exact behavior of the optimal solution with respect to variations in the cost vector and the right-hand side, only when $\gamma = \frac{1}{2}$ and $\gamma = 1$. For $\frac{1}{2} < \gamma < 1$ conservative approximations will be obtained; that is to say the optimal solutions will belong to a subset of the feasibility set and will be optimal among that subset. The motivation for using this approximation is the same as discussed previously for Model I.

C. The Probability Vector γ

γ is given a priori in Models I and II. One might be interested in the effect on the optimal value of the objective function of relaxing the probability constraints by decreasing γ or tightening the probability constraints by increasing γ. For example, if the probability constraints represent production standards, what will be the effect of modifying these standards on the optimal expected cost? The analysis can be performed in many different ways: modifying one constraint at the time, or many constraints.

1. Model I:

Both the deterministic equivalent and the conservative approximation are linear programs with γ in the right-hand side and the usual sensitivity analysis of the right-hand side can be used. The motivation for using the approximation has been discussed in the previous section. Here we may add that the sensitivity analysis can be performed easily on the approximation since the right side is a linear function of γ. If we choose to change only one γ_i or modify all γ_i $i = 1, \ldots, m$ by the same ratio, we have the usual one dimensional sensitivity analysis. However if we choose to modify more than one or in general modify γ, then this is a multidimensional sensitivity analysis and it is discussed in Walters [14].

The conservative approximation for Model I when $e_j \in C$ $j = 1, \ldots, n$ and $\gamma \geq \frac{1}{2}$ can be formulated as:

$$\text{minimize} \quad c^T x$$

$$\text{subject to}$$

$$\sum_{j=1}^{n} A_{ij} x_j \geq b_i + (2\gamma_i - 1) \sum_{j=1}^{n} |A_{ij}| a_j \qquad i = 1, \ldots, m$$

or in a more convenient notation:

$$\text{minimize} \quad c^T x$$

$$\text{subject to} \quad Ax \geq b + \tilde{B}\tilde{\gamma}$$

where B is an $n \times m$ matrix with

$$\tilde{B}_{ij} = 0 \quad \text{for} \quad i \neq j$$

and

$$\tilde{B}_{ii} = \sum_{j=1}^{n} |A_{ij}| a_j \qquad i = 1, \ldots, m$$

and

$$\tilde{\gamma}_i = 2\gamma_i - 1 \; .$$

The most general sensitivity analysis would be to find all optimal solutions for $0 \leq \tilde{\gamma} \leq 1$.

Of course if we know the exact form of the deterministic equivalent, we should perform the sensitivity analysis on this problem and get exact results. However, it cannot be done directly as sensitivity analysis of the right-hand side of a linear program, because the right-hand side is not a linear function of γ . The usual sensitivity analysis can be performed for the parameter

$\lambda = (\lambda_1, \ldots, \lambda_m)$ where $\lambda_i = G_i^{-1}(1 - \gamma_i)$ $i = 1, \ldots, m$ and then later converted back to γ using the formulas:

$\gamma_i = 1 - G_i(\lambda_i)$ $i = 1, \ldots, m$.

Here again, as described for the approximation, one-dimensional or multidimensional analysis can be performed.

2. Model II:

In this case the γ vector is not represented in the right-hand but in the left-hand side of the constraints. We will not deal with the deterministic equivalent as the left-hand side is dependent on the distribution of the random variables and is different for every problem. However we will mention that for the conservative approximation, the one-dimensional case is a parametric column linear program and this problem has been treated in Lawrence [9].

D. Variations in the Distribution Functions of the Random Variables

In this type of sensitivity analysis the probability vector γ has been specified and cannot be changed; however, the distribution functions of the random variables can be modified. For example, if the random variables represent error in measurement, replacing the measuring device with a more precise one will alter the distribution functions of the errors.

In the deterministic equivalent approach, a change in the distributions of the random variables produces a corresponding change in the convolution. Since it is very difficult to describe mathematically this relationship for a general enough case, we will concentrate our efforts on the conservative approximation approach.

1. The Change in Distribution Functions for Random Variables Belonging to C :

From the formulation of the approximations, the main relevant change in the distributions of which we want to keep track is the change in the range of the densities. However, it is important that the modified random variables still belong to C for the approximation to be meaningful. We will give two examples of such changes. A random variable $X \in C$ with density f, distribution function F and range $[-a,+a]$ will be considered. We want to reduce the range to $[-\theta a,+\theta a]$ where $0 < \theta < 1$.

Example 1: Truncation

The new random variable $\tilde{X}$ will have range $[-\theta a,+\theta a]$ and the following density $\tilde{f}$ and distribution function $\tilde{F}$.

$$\tilde{f}(x) = \begin{cases} 0 & x < -\theta a \\ \dfrac{f(x)}{F(\theta a) - F(-\theta a)} & -\theta a \leq x \leq \theta a \\ 0 & x > \theta a\ . \end{cases}$$

$$\tilde{F}(x) = \begin{cases} 0 & x < -\theta a \\ \dfrac{F(x) - F(-\theta a)}{F(\theta a) - F(-\theta a)} & -\theta a \leq x \leq \theta a \\ 1 & x \geq \theta a\ . \end{cases}$$

The properties of X are presented as shown in the following theorem:

Theorem 4.1:

If $X \varepsilon C$, then $\tilde{X} \varepsilon C$.

Proof:

a) $\tilde{f}(x) = \tilde{f}(-x) \quad \forall x$ since $f(x) = f(-x) \quad \forall x$.

b) $\tilde{f}(x)$ is unimodal with maximum at 0 since $f(x)$ is unimodal with maximum at 0.

c) $[-\theta a, +\theta a]$ is finite and is the range of $\tilde{X}$.

d) $\tilde{f}(x)$ is continuous since $f(x)$ is continuous.

Example 2: Concentration of Mass

The new random variable $\bar{X}$ will have range $[-\theta a, +\theta a]$ and the following density $\bar{f}$ and distribution function $\bar{F}$.

$$\bar{f}(x) = \begin{cases} 0 & x < -\theta a \\ \frac{1}{\theta} f\left(\frac{x}{\theta}\right) & -\theta a \leq x \leq \theta a \\ 0 & x > \theta a \ . \end{cases}$$

$$\bar{F}(x) = \begin{cases} 0 & x < -\theta a \\ F\left(\frac{x}{\theta}\right) & -\theta a \leq x \leq \theta a \\ 1 & x \geq \theta a \ . \end{cases}$$

The properties of X are preserved as is shown in Theorem 4.2.

Theorem 4.2:

If $X \varepsilon C$, $\bar{X} \varepsilon C$.

Proof:

a) For $x \varepsilon [-\theta a, +\theta a]$ $\bar{f}(x) = \frac{1}{\theta} f\left(\frac{x}{\theta}\right) = \frac{1}{\theta} f\left(\frac{-x}{\theta}\right) = \bar{f}(-x)$ as $X \varepsilon C$.

b) As $X \varepsilon C$ $f(x)$ is unimodal with maximum at 0, so is $f\left(\frac{x}{\theta}\right)$ and $\frac{1}{\theta} f\left(\frac{x}{\theta}\right)$.

c) $[-\theta a, +\theta a]$ is finite.

d) $\bar{f}(x)$ is continuous since $f(x)$ is continuous.■

This second method gives immediately the corresponding chance on the variance of the random variable.

Theorem 4.3:

For $X \varepsilon C$ and $0 < \theta < 1$

$$\mathrm{Var}\ (\bar{X}) = \theta^2\ \mathrm{Var}\ (X)\ .$$

Proof:

$$\mathrm{Var}\ (\bar{X}) = \int_{-\theta a}^{+\theta a} x^2 \bar{f}(x)dx = \int_{-\theta a}^{+\theta a} \frac{x^2}{\theta} f\left(\frac{x}{\theta}\right)dx$$

change of variables

$$v = \frac{x}{\theta}$$

$$\Longleftrightarrow x = \theta v$$

$$dx = \theta dv$$

$$\Rightarrow \text{Var }(\bar{X}) = \int_{-a}^{+a} \frac{(\theta v)^2}{\theta} f(v)\theta dv$$

$$= \theta^2 \int_{-a}^{+a} v^2 f(v)dv$$

$$= \theta^2 \text{ Var }(X) \;. \blacksquare$$

2. The Parametric Models:

θ is now the parameter for this sensitivity analysis with each random variable having its parameter θ_j $j = 1, \ldots, n$. The conservative approximation will be presented here.

Model I:

$\left(\text{For } e_j \in C \;\; j = 1, \ldots, n \text{ and } \gamma \geq \frac{1}{2}\right)$.

$$\text{minimize} \quad c^T x$$

$$\text{subject to}$$

$$\sum_{j=1}^{n} A_{ij}x_j \geq b_i + (2\gamma_i - 1) \sum_{j=1}^{n} |A_{ij}| a_j \theta_j \qquad i = 1, \ldots, m \;.$$

Using a more convenient notation the model can be described as:

$$\text{minimize} \quad c^T x$$

$$\text{subject to} \quad Ax \geq b + B'\theta$$

where B' is an $n \times m$ matrix such that $B'_{ij} = (2\gamma_i - 1)|A_{ij}|a_j$.

This is a multiparametric right-hand side linear program and such problems have been discussed in D. Walters [14]. However, any dimensional parametric problem can be discussed by using the relation $\theta = A + B\theta'$, in particular if one sets $\theta_j = \theta \ \forall j$, then it is reduced to a one-dimensional parametric problem.

Model II:

$$\left(\text{For } \alpha_j \in C \quad j = 1, \ldots, n \text{ and } \gamma \geq \frac{1}{2} \text{ and } \{x \mid x \geq 0\} \subset \Omega\right).$$

$$\text{minimize} \quad c^T x$$

$$\text{subject to} \quad x \in \Omega$$

$$\sum_{j=1}^{n} [A_{ij} + (1 - 2\gamma_i)|A_{ij}|a_j\theta_j]x_j \geq b_i \qquad i = 1, \ldots, m .$$

This is a parametric column linear program and we will refer the reader to Lawrence [9].

The deterministic equivalents for Models I and II are not in form amenable for this type of sensitivity analysis; this is why the approximations are used here. We will mention again that these approximation problems are identical to the original problems for $\gamma = 1$ and $\gamma = \frac{1}{2}$ and that these are good approximations in the neighborhood of $\gamma = 1$, as a variation θ is felt most strongly at $\gamma = 1$ as shown in Figure 4.1.

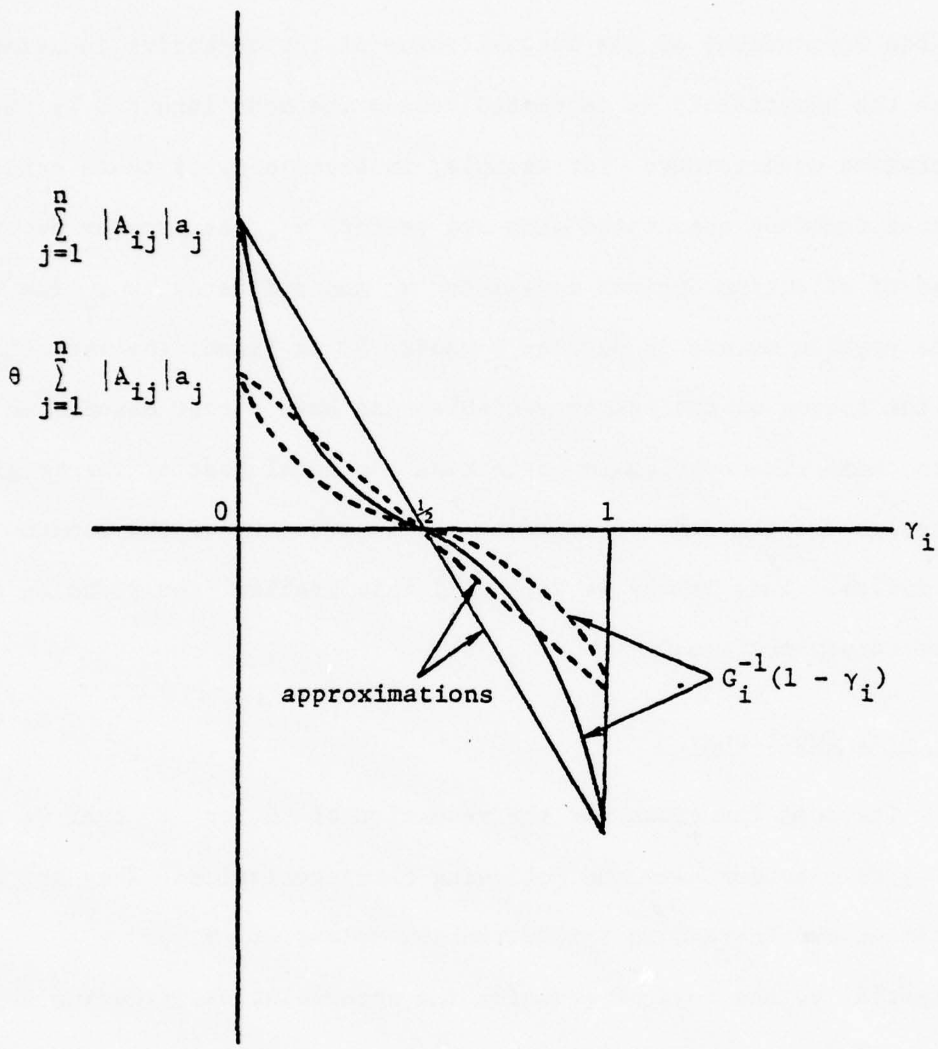

$\sum_{j=1}^{n} |A_{ij}| a_j$
$\theta \sum_{j=1}^{n} |A_{ij}| a_j$
0
½
1
γ_i
approximations
$G_i^{-1}(1 - \gamma_i)$

FIGURE 4.1

E. The Reducing of Uncertainty Problem

One problem of further interest is the problem of the trade-off of the improvement of the optimal value of the objective function when the uncertainty is increased versus the cost incurred by this increased uncertainty. For example, in Section C, if there exists a cost function associated with the vector γ , the problem become that of selecting optimal decisions x and standards γ . The same problem exists in Section D where γ is fixed; the variations in the ranges of the random variables can have a cost associated with them. The problem is optimizing the total cost of the original function and the cost of reducing the uncertainty of the random variables. This is why we will call this problem "the reducing of uncertainty problem."

1. Cost Functions:

The cost functions for the reduction of θ or γ that we are going to consider have the following characteristics: They are non-negative and increasing for decreasing values of θ or γ . Essentially, the cost of reducing the uncertainties (reducing the range $[-a_j,+a_j]$ to $[-\theta a_j,\theta a_j]$ with $0 < \theta < 1$) goes up as the range decreases. For γ , the cost due to producing material not meeting standards or not meeting the demand, for example, increases as the probability γ is decreased.

We will propose the following cost functions. We will use the symbol δ to represent both the vector θ and $\tilde{\gamma}$ ($\tilde{\gamma} = 2\gamma - 1$ as defined in 2. of Section C) to avoid repeating the discussion.

a) *In the case where all reductions* δ_i $i = 1, \ldots, n$ are equal to δ, a reasonable cost function is:

$$c(\delta) = \frac{d_0}{\delta} - d_0$$

with $d_0 \geq 0$. Then

$$c(1) = 0$$

$$\lim_{\delta \to 0} [c(\delta)] = +\infty .$$

This is nonnegative and increasing for decreasing values of δ $(0 \leq \delta \leq 1)$.

b) In the case a) above it might be unrealistic to postulate that all uncertainty can only be removed at an infinite cost. This is especially true for $\delta = \tilde{\gamma}$. Recall that $\tilde{\gamma} = 2\gamma - 1$ in the parametric problem where γ is the parameter. We can see that if the cost is infinite for $\tilde{\gamma} = 0$, this corresponds to infinite cost for $\gamma = \frac{1}{2}$, which is the cost of each constraint to be satisfied $\frac{1}{2}$ of the time. It is clear that a modification to a) is needed to reflect this fact, this is why we introduce b): If $\delta_i = \delta \quad \forall n$

$$c(\delta) = \frac{d_0}{d_1 + \delta} - \frac{d_0}{d_1 + 1} \quad \text{where } d_1 > 0 \; ; \; d_0 \geq 0$$

then $c(1) = 0 \; ; \; c(0) = \dfrac{d_0}{d_1(d_1 + 1)} < +\infty$.

c) When all γ_i are not necessarily equal, the generalization, for $\delta = (\delta_1, \ldots, \delta_n)$, is:

$$c(\delta) = \sum_{j=1}^{n} \frac{d_0^j}{d_1^j + \delta_j} - \sum_{j=1}^{n} \frac{d_0^j}{d_1^j + 1}$$

with $d_1^j \geq 0$ and $d_0^j \geq 0$ for all $j = 1, \ldots, n$.

d) A more global approach where the total uncertainty reduced is taken into consideration is for $\delta = (\delta_1, \ldots, \delta_n)$:

$$c(\delta) = \frac{1}{\sum_{j=1}^{n} d_j\delta_j + d_0} - \frac{1}{\sum_{j=0}^{n} d_j}$$

with $d_j \geq 0 \quad \forall j = 0, \ldots, n$.

It is important to note that for $\delta_j = \theta_j$, we are talking about reducing uncertainty when we reduce δ . It is in this context that the name "the reducing of uncertainty problem" was devised. However, when $\delta_j = \tilde{\gamma}_j = 2\gamma_j - 1$, a decrease in δ corresponds to a decrease in γ_i which is an increase in the probability of the solution not being feasible and the cost function then reflects the cost associated with this expected loss of feasibility.

Many more cost functions could be proposed; however, we will restrict our attention to d) which is a generalization of a) and b). This is because this cost function has the advantage of being convex as will be shown in the following theorem. It is also amenable to a solution scheme presented in Chapter V.

Theorem 4.4:

For $d_j \geq 0 \quad j = 0,1, \ldots, n$ and $y \geq 0$ the function $\frac{1}{d_0 + d^T y}$ is convex.

Proof:

Consider any vector x and y such that $x \geq 0$ and $y \geq 0$.

Then:

Case 1:

If

$$d^T x \geq d^T y \geq 0$$

$$\Rightarrow d_0 + d^T x \geq d_0 + d^T y \geq 0$$

$$\Rightarrow \frac{d^T x - d^T y}{(d_0 + d^T x)\left(d_0 + d^T y\right)} \geq \frac{d^T x - d^T y}{\left(d_0 + d^T x\right)^2} \geq 0 \ .$$

Case 2:

If

$$d^T y \geq d^T x \geq 0$$

$$\Rightarrow 0 \geq \frac{d^T x - d^T y}{\left(d_0 + d^T x\right)\left(d_0 + d^T y\right)} \geq \frac{d^T x - d^T y}{\left(d_0 + d^T x\right)^2} \ .$$

From Case 1 and 2 it follows that

$$\frac{1}{d_0 + d^T y} - \frac{1}{d_0 + d^T x} \geq \frac{d^T x - d^T y}{\left(d_0 + d^T x\right)^2}$$

$$\Rightarrow \frac{1}{d_0 + d^T y} \geq \frac{1}{d_0 + d^T x} - \sum_{j=1}^{n} \frac{d_j (y_j - x_j)}{\left(d_0 + d^T x\right)^2}$$

$$\Longleftrightarrow \frac{1}{d_0 + d^T y} \geq \frac{1}{d_0 + d^T x} + \nabla\left[\frac{1}{d_0 + d^T x}\right] \times [y - x] \ .$$

Therefore $\dfrac{1}{d_0 + d^T x}$ is a convex function. ■

2. Applications to Model I:

For $\gamma \geq \frac{1}{2}$ and $e_j \in C \quad j = 1, \ldots, n$ we have the following conservative approximations.

a) For $\delta = \tilde{\gamma}$.

The model can be formulated as follows using the notation of 1) of Section 3.

$$\text{minimize} \quad c^T x + \frac{1}{d_0 + d^T \tilde{\gamma}} - \frac{1}{\sum_{j=0}^{n} d_j}$$

$$\text{subject to} \quad Ax - \tilde{B}\tilde{\gamma} \geq b$$

$$1 \geq \tilde{\gamma} \geq 0 \ .$$

b) For $\delta = \theta$.

Here also using the notation 2) of Section D the model is

$$\text{minimize} \quad c^T x + \frac{1}{d_0 + d^T\theta} - \frac{1}{\sum_{j=0}^{n} d_j}$$

$$\text{subject to} \quad Ax - B'\theta \geq b$$

$$1 \geq \theta \geq 0 \; .$$

These two problems a) and b) fall into the general pattern of the following problem:

$$\text{minimize} \quad c^T x + \frac{1}{d_0 + d^T y} - \frac{1}{\sum_{j=0}^{n} d_j}$$

$$\text{subject to} \quad Ax - By \geq b$$

$$1 \geq y \geq 0$$

with $d \geq 0$.

This problem is analyzed in Chapter V and an algorithm is presented to solve it.

These conservative approximations do not yield the exact solution. However, these problems are exact for $\gamma = 1$ or $\frac{1}{2}$, and for b) the approximation is very close for γ near 1 as the variable θ reflects the change in the ranges of the random variables and this variation is felt most strongly in the neighborhood of $\gamma = 1$ as mentioned earlier (see Figure 4.1).

3. Applications to Model II:

For $\gamma \geq \frac{1}{2}$ and $\alpha_j \in C$ $j = 1, \ldots, n$ we have the following conservative approximations.

a) For $\delta = \gamma$

$$\text{minimize} \quad c^T x + \frac{1}{d_0 + d^T \tilde{\gamma}} - \frac{1}{\sum_{j=0}^{n} d_j}$$

subject to

$$\sum_{j=1}^{n} [A_{ij} - \tilde{\gamma} |A_{ij}| a_j] x_j \geq b_i \qquad i = 1, \ldots, m$$

$$x \geq 0$$

$$1 \geq \tilde{\gamma} \geq 0 .$$

b) For $\delta = \theta$

$$\text{minimize} \quad c^T x + \frac{1}{d_0 + d^T \theta} - \frac{1}{\sum_{j=1}^{n} d_j}$$

subject to

$$\sum_{j=1}^{n} [A_{ij} + [1 - 2\gamma_i] |A_{ij}| a_j \theta_j] x_j \geq b_i \qquad i = 1, \ldots, m$$

$$x \geq 0$$

$$1 \geq \theta \geq 0 .$$

These two problems are nonlinear programs that can be solved using the known algorithm and no special solution schemes will be presented.

CHAPTER V

In this chapter the following programming problem is analyzed and an algorithm is specified for its solution.

A. The Reducing of Uncertainty Problem

Find the optimal solution (x,y) which solves

$$\text{minimize} \quad c^T x + \frac{1}{d_0 + d^T y}$$

$$\text{subject to} \quad Ax - By \geq b$$

$$x \geq 0$$

$$1 \geq y \geq 0$$

with $d \geq 0$ and $d_0 \geq 0$.

This is a convex program as the feasibility set is a polyhedral set and therefore convex, and its objective function is convex as the sum of two convex functions since $c^T x$ is linear and $\frac{1}{d_0 + d^T y}$ is convex as shown in Theorem 4.4.

This problem is clearly defined if the following condition holds:

Condition 5.1:

There exists $(\bar{x},\bar{y})$ such that

$$A\bar{x} - B\bar{y} \geq b$$

$$\bar{x} \geq 0$$

$$1 \geq \bar{y} \geq 0$$

and $d_0 + d^T\bar{y} > 0$.

We will assume this condition is satisfied throughout this discussion.

The strategy used to solve this problem will consist of solving a different equivalent program and obtaining the original optimal solution from the optimal solution of the second program. The term equivalent program is used as defined below.

Definition:

Two programs are equivalent if there exists a scheme to obtain the optimal solution of any one of the two problems knowing the optimal solution of the other problem.

B. Two Equivalent Convex Programs

Theorem 5.2:

The following two programs are equivalent:

(a) $\underset{x\varepsilon\Omega}{\text{minimize}} \; h(x)$

where $h(x) = f(x) + g(x)$, $f(x)$ and $g(x)$ are convex functions for $x \varepsilon \Omega$, and Ω is a convex set.

(b) $\underset{\alpha\varepsilon\theta}{\text{minimize}} \; \phi(\alpha)$

where

$$\phi(\alpha) = \alpha + \min_{x\varepsilon\Omega} \{g(x) \mid f(x) \leq \alpha\}$$

and

$$\theta = \{\alpha \mid \{x \mid x \in \Omega \text{ and } f(x) \leq \alpha\} \neq \emptyset\} .$$

In other words, using the definition of equivalent programs, the following statements are true:

(i) $\min_{x\in\Omega} h(x) = \min_{\alpha\in\theta} \phi(\alpha)$.

(ii) For any $\bar{\alpha}$ optimal solution of (b) there exists at least one corresponding $\bar{x}$ optimal solution of (a), which can be obtained as follows: $\bar{x} \in \Omega$ such that $\min_{x\in\Omega} \{g(x) \mid f(x) \leq \bar{\alpha}\} = g(\bar{x})$.

(iii) For any $\bar{x}$ optimal solution of (a) there exists at least one corresponding $\bar{\alpha}$ optimal solution of (b), which can be obtained as follows: $\bar{\alpha}$ such that $\bar{\alpha} = f(\bar{x})$ and $\bar{\alpha} \in \theta$.

Proof:

Let us first establish two facts to be used later in the proof: For any $\bar{x}$ such that $\bar{x} \in \Omega$ and $h(\bar{x}) = \min_{x\in\Omega} h(x)$ then:

Fact 1:

$$g(\bar{x}) = \min_{x\in\Omega} \{g(x) \mid f(x) \leq f(\bar{x})\} .$$

Fact 2:

$$\phi(f(\bar{x})) \leq \min_{\alpha\in\theta} \phi(\alpha) .$$

If Fact 1 were not true then there would exist an $\bar{\bar{x}}$ such that $\bar{\bar{x}} \varepsilon \Omega$ and $g(\bar{\bar{x}}) < g(\bar{x})$ and $f(\bar{\bar{x}}) \leq f(\bar{x})$ which yields: $\bar{\bar{x}} \varepsilon \Omega$ such that $h(\bar{\bar{x}}) = f(\bar{\bar{x}}) + g(\bar{\bar{x}}) < f(\bar{x}) + g(\bar{x}) = h(\bar{x})$.

This contradicts our assumption about $\bar{x}$.

Therefore Fact 1 is true.

If Fact 2 were not true then there would exist an $\bar{\alpha}$ such that $\bar{\alpha} \varepsilon \theta$ and $\phi(\bar{\alpha}) < \phi(f(\bar{x}))$ as $\phi(\bar{\alpha}) = \bar{\alpha} + \min\limits_{x \varepsilon \Omega} \{g(x) \mid f(x) \leq \bar{\alpha}\}$ and $\bar{\alpha} \varepsilon \theta$. There exists an x^* such that $x^* \varepsilon \Omega$ and

$$g(x^*) = \min_{x \varepsilon \Omega} \{g(x) \mid f(x) \leq \bar{\alpha}\} .$$

Therefore, there exists $x^* \varepsilon \Omega$ such that $h(x^*) = f(x^*) + g(x^*) \leq \bar{\alpha} + g(x^*) = \phi(\bar{\alpha})$.

Hence

$$h(x^*) < \phi(f(\bar{x})) .$$

However, Fact 1 established that

$$\phi(f(\bar{x})) = f(\bar{x}) + \min_{x \varepsilon \Omega} \{g(x) \mid f(x) \leq f(\bar{x})\} = f(\bar{x}) + g(\bar{x})$$

so there exists $x^* \varepsilon \Omega$ such that

$$h(x^*) < f(\bar{x}) + g(\bar{x}) = h(\bar{x})$$

which contradicts the assumption about $\bar{x}$, therefore, Fact 2 is true.

Now we can prove:

(i) $\min_{x \in \Omega} h(x) = \min_{\alpha \in \theta} \phi(\alpha)$.

As Fact 2 say that

$$\phi(f(\bar{x})) \leq \min_{\alpha \in \theta} \phi(\alpha) .$$

It is sufficient to show that:

1) $f(\bar{x}) \in \theta$.

This is true since

$$\bar{x} \in \{x \mid x \in \Omega \text{ and } f(x) \leq f(\bar{x})\}$$

and therefore $\{x \mid x \in \Omega \text{ and } f(x) \leq f(\bar{x})\} \neq \emptyset$.

2) $h(\bar{x}) = \phi(f(\bar{x}))$.

Fact 1 establishes this as:

$$\begin{aligned}\phi(f(\bar{x})) &= f(\bar{x}) + \min_{x \in \Omega} \{g(x) \mid f(x) \leq f(\bar{x})\} \\ &= f(\bar{x}) + g(\bar{x}) \\ &= h(\bar{x}) .\end{aligned}$$

Hence as $f(\bar{x}) \in \theta$ and by Fact 2

$$\phi(f(\bar{x})) = \min_{\alpha \in \theta} \phi(\alpha)$$

and since

$$\phi(f(\bar{x})) = h(\bar{x})$$

and

$$h(\bar{x}) = \min_{x \varepsilon \Omega} h(x) \ .$$

It is true that

$$\min_{x \varepsilon \Omega} h(x) = \min_{\alpha \varepsilon \theta} \phi(\alpha) \ .$$

(ii) $\forall \bar{\alpha} \ \varepsilon \ \theta$ such that

$$\phi(\bar{\alpha}) = \min_{\alpha \varepsilon \theta} \phi(\alpha) \ .$$

There exists $\bar{x} \ \varepsilon \ \Omega$ such that

$$g(\bar{x}) = \min_{x \varepsilon \Omega} \{g(x) \mid f(x) \le \bar{\alpha}\}$$

and

$$h(\bar{x}) = f(\bar{x}) + g(\bar{x}) = \min_{x \varepsilon \Omega} h(x) \ .$$

Let $\bar{x} \ \varepsilon \ \Omega$ be such that $g(\bar{x}) = \min_{x \varepsilon \Omega} \{g(x) \mid f(x) \le \bar{\alpha}\}$; such an $\bar{x}$ exists since $\bar{\alpha} \ \varepsilon \ \theta$, then

$$f(\bar{x}) = \bar{\alpha} \quad \text{and} \quad h(\bar{x}) = \phi(\bar{\alpha})$$

because if it were not true then $f(\bar{x}) < \bar{\alpha}$ and

$$h(\bar{x}) = f(\bar{x}) + g(\bar{x}) < \phi(\bar{\alpha})$$

which yields the contradiction that there exists $\bar{x} \in \Omega$ such that

$$h(\bar{x}) < \min_{\alpha \in \theta} \phi(\alpha) = \min_{x \in \Omega} h(x)$$

as seen in (i). Therefore

$$\phi(\bar{\alpha}) = \bar{\alpha} + g(\bar{x}) = f(\bar{x}) + g(\bar{x}) = h(\bar{x}) \ .$$

Hence $h(\bar{x}) = \min_{\alpha \in \theta} \phi(\alpha) = \min_{x \in \Omega} h(x)$ as seen in (i).

(iii) $\forall \bar{x}$ such that $\bar{x} \in \Omega$ and $h(\bar{x}) = \min_{x \in \Omega} h(x)$, then there exists $\bar{\alpha} \in \theta$ such that $\bar{\alpha} = f(\bar{x})$ and

$$\phi(\bar{\alpha}) = \min_{\alpha \in \theta} \phi(\alpha) \ .$$

Let $\bar{\alpha} \in \theta$ such that $\bar{\alpha} = f(\bar{x})$; such $\bar{\alpha}$ exists since $\{x \mid x \in \Omega \text{ and } f(x) \leq f(\bar{x})\} \neq \emptyset$ since $\bar{x}$ belongs to that set.

We have seen in (i) and (ii) that

$$h(\bar{x}) = \min_{\alpha \in \theta} \phi(\alpha)$$

and

$$h(\bar{x}) = \phi(f(\bar{x})) \ .$$

Hence

$$\phi(f(\bar{x})) = \min_{\alpha \in \theta} \phi(\alpha)$$

and $\bar{\alpha} = f(\bar{x})$ is an optimal solution of (b).■

Program (b) is also a convex program as is seen in the next two lemmas.

Lemma 5.3:

The set $\theta = \{\alpha \mid \{x \mid x \in \Omega \text{ and } f(x) \leq \alpha\} \neq \emptyset\}$ is a convex set.

Proof:

$\forall \alpha^1 \in \theta$ and $\alpha^2 \in \theta$. There exist $x^1 \in \Omega$ and $x^2 \in \Omega$ such that

$$f(x^1) \leq \alpha^1 \text{ and } f(x^2) \leq \alpha^2 .$$

Hence for any λ such that $0 \leq \lambda \leq 1$ it is true that:

$$\lambda f(x^1) + (1 - \lambda) f(x^2) \leq \lambda\alpha^1 + (1 - \lambda)\alpha^2 .$$

However using the convexity of $f(x)$ it is also true that:

$$f(\lambda x^1 + (1 - \lambda)x^2) \leq \lambda f(x^1) + (1 - \lambda) f(x^2) \leq \lambda\alpha^1 + (1 - \lambda)\alpha^2 .$$

This inequality and the fact that Ω is a convex set clearly shows that for any $\alpha^1 \in \theta$, $\alpha^2 \in \theta$, $0 \leq \lambda \leq 1$ there exists an $x \in \Omega$, namely $\lambda x^1 + (1 - \lambda)x^2$, such that $f(x) \leq \lambda\alpha^1 + (1 - \lambda)\alpha^2$.

Therefore $\lambda\alpha^1 + (1 - \lambda)\alpha^2 \in \theta$ and θ is a convex set. ■

Lemma 5.4:

The function $\phi(\alpha)$ is convex for $\alpha \in \theta$.

Proof:

$\forall \alpha^1 \varepsilon \theta$, $\alpha^2 \varepsilon \theta$.

$$\phi(\alpha^1) = \alpha^1 + \min_{x\varepsilon\Omega} \{g(x) \mid f(x) \leq \alpha^1\} .$$

$$\phi(\alpha^2) = \alpha^2 + \min_{x\varepsilon\Omega} \{g(x) \mid f(x) \leq \alpha^2\} .$$

As $\alpha^1 \varepsilon \theta$ and $\alpha^2 \varepsilon \theta$ there exists $x^1 \varepsilon \Omega$ and $x^2 \varepsilon \Omega$ such that

$$\phi(\alpha^1) = \alpha^1 + g(x^1) \quad \text{with} \quad f(x^1) \leq \alpha^1 .$$

$$\phi(\alpha^2) = \alpha^2 + g(x^2) \quad \text{with} \quad f(x^2) \leq \alpha^2 .$$

For any λ such that $0 \leq \lambda \leq 1$, $\lambda\alpha^1 + (1 - \lambda)\alpha^2 \varepsilon \theta$ by Lemma 5.3, therefore there exists:

$$\phi(\lambda\alpha^1 + (1 - \lambda)\alpha^2) = \lambda\alpha^1 + (1 - \lambda)\alpha^2$$

$$+ \min_{x\varepsilon\Omega} \{g(x) \mid f(x) \leq \lambda\alpha^1 + (1 - \lambda)\alpha^2\} .$$

Now if we want to prove that $\phi(\alpha)$ is a convex function for $\alpha \varepsilon \theta$ it is sufficient to show that

$$\phi(\lambda\alpha^1 + (1 - \lambda)\alpha^2) \leq \lambda\phi(\alpha^1) + (1 - \lambda)\phi(\alpha^2) .$$

Since f is a convex function and Ω a convex set

$$\lambda x^1 + (1 - \lambda)x^2 \varepsilon \Omega$$

and

$$f(\lambda x^1 + (1 - \lambda)x^2) \le \lambda f(x^1) + (1 - \lambda)f(x^2) \le \lambda\alpha^1 + (1 - \lambda)\alpha^2 .$$

Therefore:

$$\min_{x\varepsilon\Omega} \{g(x) \mid f(x) \le \lambda\alpha^1 + (1 - \lambda)\alpha^2\} \le g(\lambda x^1 + (1 - \lambda)x^2) .$$

Hence

$$\phi(\lambda\alpha^1 + (1 - \lambda)\alpha^2) \le \lambda\alpha^1 + (1 - \lambda)\alpha^2 + g(\lambda x^1 + (1 - \lambda)x^2) .$$

Using now the convexity of $g(x)$ it is true that:

$$\phi(\lambda\alpha^1 + (1 - \lambda)\alpha^2) \le \lambda\alpha^1 + (1 - \lambda)\alpha^2 + \lambda g(x^1) + (1 - \lambda)g(x^2)$$

$$\Longleftrightarrow \phi(\lambda\alpha^1 + (1 - \lambda)\alpha^2) \le \lambda[\alpha^1 + g(x^1)] + (1 - \lambda)[\alpha^2 + g(x^2)]$$

$$\Longleftrightarrow \phi(\lambda\alpha^1 + (1 - \lambda)\alpha^2) \le \lambda\phi(\alpha^1) + (1 - \lambda)\phi(\alpha^2)$$

$\phi(\alpha)$ is convex for $\alpha \varepsilon \theta$. ■

C. An Equivalent Convex Program to the Reducing of Uncertainty Problem

If we apply the results of Section B to our original problem we obtain the following facts:

Theorem 5.5:

The following two problems are equivalent.

(a) minimize $c^T x + \dfrac{1}{d_0 + d^T y}$

subject to $Ax - By \geq b$

$x \geq 0$

$1 \geq y \geq 0$.

(b) minimize $\phi(\alpha)$

subject to $\alpha \in \theta$

where

$$\phi(\alpha) = \alpha + \min_{(x,y)\in\Omega} \left\{ \frac{1}{d_0 + d^T y} \mid c^T x \leq \alpha \right\}$$

$$\Omega = \{(x,y) \mid Ax - By \geq b \ , \ x \geq 0 \ , \ 1 \geq y \geq 0\}$$

$$\theta = \{\alpha \mid \{(x,y) \mid (x,y) \in \Omega \text{ and } c^T x \leq \alpha\} \neq \emptyset\} \ .$$

Proof:

Apply Theorem 5.2 to the convex program (a).■

Since it is our goal to use problem (b) to solve problem (a) it is useful to study (b) further.

Proposition 5.6:

Under condition 5.1:

$$\min_{(x,y)\in\Omega} \left\{ \frac{1}{d_0 + d^T y} \mid c^T x \leq \alpha \right\} = \frac{-1}{\min\limits_{(x,y)\in\Omega} \left\{ -d_0 - d^T y \mid c^T x \leq \alpha \right\}} \ .$$

Proof:

Since $d_0 \geq 0$ and $d \geq 0$ and $x \geq 0$, $y \geq 0$

$$\min_{(x,y)\varepsilon\Omega} \left\{ \frac{1}{d_0 + d^T y} \mid c^T x \le \alpha \right\} = \frac{1}{\max\limits_{(x,y)\varepsilon\Omega} \left\{ d_0 + d^T y \mid c^T x \le \alpha \right\}}$$

$$= \frac{1}{-\min\limits_{(x,y)\varepsilon\Omega} \left\{ -d_0 - d^T y \mid c^T x \le \alpha \right\}} = \frac{-1}{\min\limits_{(x,y)\varepsilon\Omega} \left\{ -d_0 - d^T y \mid c^T x \le \alpha \right\}} . \blacksquare$$

Corollary 5.7:

$$\phi(\alpha) = \alpha + \frac{1}{\psi(\alpha)}$$

where

$$\psi(\alpha) = -\min_{(x,y)\varepsilon\Omega} \left\{ -d_0 - d^T y \mid c^T x \le \alpha \right\} .$$

Proof:

Apply Proposition 5.6.$\blacksquare$

Proposition 5.8:

$\psi(\alpha)$ is a bounded function for $\alpha \in \theta$. $\sum_{j=0}^{n} d_j \ge \psi(\alpha) \ge 0$.

Proof:

Since $d_0 \ge 0$ and $d \ge 0$ and as $(x,y) \in \Omega \Rightarrow 1 \ge y \ge 0 \Rightarrow$

$0 \le d_0 + d^T y \le \sum_{j=0}^{n} d_j \Rightarrow + \sum_{j=0}^{n} d_j \ge \psi(\alpha) \ge 0$.$\blacksquare$

Proposition 5.9:

(i) If for $\bar{\alpha}$ $\psi(\bar{\alpha}) = \sum_{j=0}^{n} d_j$. Then $\forall\alpha$ subject to $\alpha \ge \bar{\alpha}$

$$\psi(\alpha) = \sum_{j=0}^{n} d_j .$$

(ii) If for $\bar{\alpha}$ $\psi(\bar{\alpha}) = 0$. Then $\forall\alpha$ subject to $\alpha \leq \bar{\alpha}$

$$\psi(\alpha) = 0 \text{ or } \psi(\alpha) \text{ is infeasible.}$$

Proof:

Since $\psi(\alpha) = - \min\limits_{(x,y)\epsilon\Omega} \left\{-d_0 - d^Ty \mid c^Tx \leq \alpha\right\}$

(i) If $\psi(\bar{\alpha}) = \sum\limits_{j=0}^{n} d_j \Rightarrow$ for $c^Tx \leq \bar{\alpha}$ the minimum bound $-\sum\limits_{j=0}^{n} d_j$ has been reached and relaxing the constraint $c^Tx \leq \alpha$ will not improve the objective function.

(ii) Same reasoning as $-d_0 - d^Ty$ reaches its upper bound; however, $\psi(\alpha)$ can be infeasible for $\alpha < \bar{\alpha}$.

Corollary 5.10:

If for $\bar{\alpha}$ $\psi(\bar{\alpha}) = \sum\limits_{j=0}^{n} d_j$. *Then* $\phi(\alpha)$ *is an increasing linear* function of α for $\alpha \geq \bar{\alpha}$: $\phi(\alpha) = \alpha + \dfrac{1}{\sum\limits_{j=0}^{n} d_j}$.

Proof:

$$\phi(\alpha) = \alpha + \frac{1}{\psi(\alpha)} .$$

Proposition 5.9(i) shows that for $\alpha \geq \bar{\alpha}$ $\psi(\alpha) = \sum\limits_{j=0}^{n} d_j$, therefore for $\alpha \geq \bar{\alpha}$ $\phi(\alpha) = \alpha + \dfrac{1}{\sum\limits_{j=0}^{n} d_j}$.■

Proposition 5.11:

The set θ is either the empty set or the whole real line or an interval closed on the left and having $+\infty$ as a boundary on the right.

<u>Proof</u>:

Lemma 5.3 established that θ is a convex set and as it is one-dimensional, θ is an interval of the real line. If the constraint set Ω is empty, θ is empty as there does not exist any (x,y) such that $(x,y) \in \Omega$ and $c^T x \leq \alpha$ for any α .

If the problem is unbounded, Corollary 5.7 tells us that

$$\phi(\alpha) = \alpha + \frac{1}{\psi(\alpha)}$$

and Proposition 5.8

$$\alpha + \frac{1}{+\sum_{j=0}^{n} d_j} \leq \phi(\alpha) .$$

So that $\phi(\alpha)$ is unbounded if and only if $\theta = (-\infty,+\infty)$.

The upper bound of the interval is always $+\infty$ as for any $\bar{\alpha} \in \theta$. Then for any $\alpha \geq \bar{\alpha}$, $\alpha \in \theta$ since

$$\theta = \{\alpha \mid \{(x,y) \mid (x,y) \in \Omega \text{ and } c^T x \leq \alpha\} \neq 0\}$$

if there exists $(\bar{x},\bar{y}) \in \Omega$ such that $c^T\bar{x} \leq \bar{\alpha}$, then surely $c^T\bar{x} \leq \alpha$ for $\alpha \geq \bar{\alpha}$.

If there exists $\alpha = b$ such that

$$\{(x,y) \mid (x,y) \in \Omega \text{ and } c^T x < b\} = \emptyset$$

and

$$\{(x,y) \mid (x,y) \in \Omega \text{ and } c^T x \leq b\} \neq \emptyset$$

then b is the left hand bound of the interval and the interval is closed because the set

$$\{(x,y) \mid (x,y) \in \Omega \text{ and } c^T x \leq \alpha\}$$

is a closed set. ■

D. General Description of the Algorithm and Optimality Criterions

As we have mentioned, the idea for solving the original problem is to first convert it into an equivalent convex program, solve this problem, then retrieve the optimal solution. The algorithm is concerned with solving the second problem. It will be done in two phases.

Phase 1:

A sequence $\{\alpha_i\}$ is generated $(\alpha_i \in \theta)$ such that $\alpha_{i+1} > \alpha_i$ from a starting feasible point $\alpha_o \in \theta$. (Incidentally i can be negative as the starting point is not necessarily at the boundary of the set θ.) This sequence is generated until an interval $[\alpha_i, \alpha_{i+1}]$ or $[\alpha_{i-1}, \alpha_{i+1}]$ is identified where the optimal solution lies. This is done simply by evaluating the $\phi(\alpha_i)$'s and comparing them until the following optimality criterion 1 is satisfied. This criterion is to be used when the possibility of an unbounded problem has been discarded.

Optimality Criterion 1:

(i) If for α , α' , $\alpha'' \varepsilon\ \theta$ such that $\alpha' < \alpha < \alpha''$ it is true that

$$\phi(\alpha) < \phi(\alpha') \text{ and } \phi(\alpha) < \phi(\alpha'') .$$

Then the optimal solution lies in the interval $[\alpha',\alpha'']$.

(ii) If for α' , $\alpha'' \varepsilon\ \theta$ such that $\alpha' < \alpha''$ it is true that

$$\phi(\alpha') = \phi(\alpha'') .$$

Then $\forall\alpha\ \varepsilon\ [\alpha',\alpha'']$, α is an optimal solution.

(iii) If for α' , $\alpha'' \varepsilon\ \theta$ such that $\alpha' < \alpha''$ it is true that $\phi(\alpha') < \phi(\alpha'')$ and $\forall\alpha\ \varepsilon\ \theta$, $\alpha \geq \alpha'$, then the optimal solution lies in the interval $[\alpha',\alpha'']$.

Proof:

As $\phi(\alpha)$ is a convex function and θ is an interval as shown in Lemma 5.4 and Proposition 5.11; Case (i) is derived from the fact that $\phi(\alpha)$ is unimodal. Case (ii) is true because a convex function can be constant only at its minimum. Case (iii) is when α' is the left-hand boundary of the interval θ ; then as $\phi(\alpha)$ is convex, either α' is the optimum or the optimum is in the interval $[\alpha',\alpha'']$.■

Phase 2:

Once such an interval has been identified, a minimum is found for the range and this is the global minimum.

Optimality Criterion 2:

If for $\bar{\alpha}$ such that $\bar{\alpha} \in [\alpha',\alpha'']$ where $[\alpha',\alpha'']$ is the interval obtained from Phase 1, it is true that

$$\phi(\bar{\alpha}) \leq \phi(\alpha) \quad \forall \alpha \in [\alpha',\alpha''] ,$$

then $\bar{\alpha}$ is the optimal solution.

Proof:

Since $\phi(\alpha)$ is convex over the interval θ and $[\alpha',\alpha'']$ was found to contain the optimal solution, this optimality criterion is just stating that a local minimum is a global minimum. ∎

E. Feasibility and Unboundness

In this algorithm an initial problem is used to obtain a starting feasible point or to determine whether the original problem is unbounded or is infeasible.

Initial Problem:

$$\begin{aligned} \text{minimize} \quad & c^T x - d^T y - d_0 \\ \text{subject to} \quad & Ax - By \geq b \\ & x \geq 0 \\ & 1 \geq y \geq 0 \; . \end{aligned}$$

Theorem 5.12:

If the initial problem is infeasible, the original problem is infeasible.

Proof:

Both problems have the same feasibility set.∎

Theorem 5.13:

If the initial problem is unbounded, then the original problem is unbounded.

Proof:

If the initial problem is unbounded, then there exists a vector $(\bar{u},\bar{w},\overline{z^1},\overline{z^2})$ such that

$$0 \neq (\bar{u},\bar{w},\overline{z^1},\overline{z^2}) \geq 0$$

and

$$A\bar{u} - B\bar{w} - I\overline{z^1} = 0$$

$$I\bar{w} + I\overline{z^2} = 0$$

and

$$c^T\bar{u} - d^T\bar{w} < 0 \ .$$

It is obvious that $\bar{w} = \overline{z^2} = 0$ since that is the unique solution of $I(\bar{w} + \overline{z^2}) = 0 \quad \bar{w} \geq 0 \ , \ \overline{z^2} \geq 0$.

Hence, we have the vector $(\bar{u},0,\overline{z^1},0)$ such that

$$0 \neq (\bar{u},0,\overline{z^1},0) \geq 0$$

and

$$A\bar{u} - I\overline{z^1} = 0$$

and

$$c^T\bar{u} < 0 \; .$$

By Condition 5.1 there exists $(\bar{x},\bar{y})$ such that $A\bar{x} - B\bar{y} \geq b$, $\bar{x} \geq 0$, $1 \geq \bar{y} \geq 0$ and $d_0 + d^T\bar{y} > 0$. Then, the vector $(\lambda\bar{u} + \bar{x},\bar{y})$ is a feasible solution and makes the original problem unbounded by increasing λ to infinity $(\lambda \geq 0)$

$$\lim_{\lambda \to +\infty} \left(c^T(\lambda\bar{u}) + \frac{1}{d_0 + d^T\bar{y}} \right) = -\infty$$

since $c^T\bar{u} < 0$ and $\dfrac{1}{d_0 + d^T\bar{y}} < +\infty$. ■

Theorem 5.14:

If there exists $(\bar{x},\bar{y}) \in \Omega$ such that

$$c^T\bar{x} - d^T\bar{y} - d_0 = \min_{(x,y)\in\Omega} \left\{ c^Tx - d^Ty - d_0 \right\}$$

where

$$\Omega = \left\{ (x,y) \;\middle|\; \begin{array}{ll} Ax - By \geq 0 & \\ x \geq 0 & 1 \geq y \geq 0 \end{array} \right\} .$$

Then

$$-d^T\bar{y} - d_0 = \min_{(x,y)\in\Omega} \left\{ -d^Ty - d_0 \mid c^Tx \leq c^T\bar{x} \right\} .$$

Proof:

If there existed $(\bar{\bar{x}},\bar{\bar{y}}) \in \Omega$ such that $c^T\bar{\bar{x}} \leq c^T\bar{x}$ and $-d^T\bar{\bar{y}} - d_0 < -d^T\bar{y} - d_0$ then $c^T\bar{\bar{x}} - d^T\bar{\bar{y}} - d_0 < c^T\bar{x} - d^T\bar{y} - d_0$ for $(\bar{\bar{x}},\bar{\bar{y}}) \in \Omega$ which contradicts our assumption. ■

Corollary 5.15:

Under the assumption of Theorem 5.14 the following is true:

$$\psi(\alpha_o) = d^T\bar{y} + d_0$$

where $\alpha_o = c^T\bar{x}$.

Proof:

Use the definition of $\psi(\alpha)$. ■

F. The Parametric Linear Program $\psi(\alpha)$

We have already introduced $\psi(\alpha)$ noting that $\phi(\alpha) = \alpha + \dfrac{1}{\psi(\alpha)}$.

Phase 1:

As $\psi(\alpha)$ is a parametric linear program the sequence $\{\alpha_i\}$ used for Phase 1 of the algorithm will come naturally from the values of α where the parametric linear program changes basis.

$$\begin{aligned} \psi(\alpha) = -\text{minimize} \quad & -d_0 - d^Ty \\ \text{subject to} \quad & Ax - By \geq b \\ & x \geq 0 \\ & 1 \geq y \geq 0 \\ & c^Tx \leq \alpha \ . \end{aligned}$$

The initial problem gives us a starting feasible point α_o with the feasible solution for $\psi(\alpha_o)$ as seen in Corollary 5.15. The idea is then to first increase α until the optimality criterion 1 is satisfied or, if necessary, decrease α .

Phase 2:

In Phase 2 we will use the fact that the points of the sequence $\{\alpha_i\}$ are the points where the basis changes in the parametric l.p. $\psi(\alpha)$.

From the theory of linear programming it is well known that for points of the sequence $\{\alpha_i\}$:

For $\alpha_i \leq \alpha \leq \alpha_{i+1}$

$$\psi(\alpha) = \frac{\psi(\alpha_{i+1}) - \psi(\alpha_i)}{\alpha_{i+1} - \alpha_i}\alpha + \frac{\alpha_{i+1}\psi(\alpha_i) - \alpha_i\psi(\alpha_{i+1})}{\alpha_{i+1} - \alpha_i}$$

and for $\alpha_{i-1} \leq \alpha \leq \alpha_i$

$$\psi(\alpha) = \frac{\psi(\alpha_i) - \psi(\alpha_{i-1})}{\alpha_i - \alpha_{i-1}}\alpha + \frac{\alpha_i\psi(\alpha_{i-1}) - \alpha_{i-1}\psi(\alpha_i)}{\alpha_i - \alpha_{i-1}} .$$

We can now prove the following two theorems under the following assumptions: α_{i-1} , α_i , α_{i+1} are 3 consecutive points of the sequence generated from the parametric linear program $\psi(\alpha)$ which satisfied the optimality criterion 1.

Theorem 5.16:

For an interval $[\alpha_i, \alpha_{i+1}]$ such that $\phi(\alpha_i) < \phi(\alpha_{i+1})$,

(i) If $\psi(\alpha_{i+1}) - \psi(\alpha_i) \leq 0$ then

$$\phi(\alpha_i) = \min_{\alpha \varepsilon [\alpha_i, \alpha_{i+1}]} \phi(\alpha) .$$

(ii) If $\psi(\alpha_{i+1}) - \psi(\alpha_i) > 0$ then one of the following is true.

(a) $\alpha^* \varepsilon [\alpha_i, \alpha_{i+1}]$ and $\phi(\alpha^*) = \min\limits_{\alpha \varepsilon [\alpha_i, \alpha_{i+1}]} \phi(\alpha)$

(b) $\alpha^{**} \varepsilon [\alpha_i, \alpha_{i+1}]$ and $\phi(\alpha^{**}) = \min\limits_{\alpha \varepsilon [\alpha_i, \alpha_{i+1}]} \phi(\alpha)$

(c) $\alpha^* \notin [\alpha_i, \alpha_{i+1}]$ and $\alpha^{**} \notin [\alpha_i, \alpha_{i+1}]$ and

$\phi(\alpha_i) = \min\limits_{\alpha \varepsilon [\alpha_i, \alpha_{i+1}]} \phi(\alpha)$

where α^* , α^{**} are defined to be:

$$\alpha^* = \frac{\alpha_i \psi(\alpha_{i+1}) - \alpha_{i+1} \psi(\alpha_i)}{\alpha_{i+1} - \alpha_i} + \sqrt{\frac{\alpha_{i+1} - \alpha_i}{\psi(\alpha_{i+1}) - \psi(\alpha_i)}}$$

$$\alpha^{**} = \frac{\alpha_i \psi(\alpha_{i+1}) - \alpha_{i+1} \psi(\alpha_i)}{\alpha_{i+1} - \alpha_i} - \sqrt{\frac{\alpha_{i+1} - \alpha_i}{\psi(\alpha_{i+1}) - \psi(\alpha_i)}} .$$

Proof:

For the interval $[\alpha_i, \alpha_{i+1}]$

$$\phi(\alpha) = \alpha + \frac{1}{\psi(\alpha)}$$

where

$$\psi(\alpha) = \frac{\psi(\alpha_{i+1}) - \psi(\alpha_i)}{\alpha_{i+1} - \alpha_i}\,\alpha + \frac{\alpha_{i+1}\psi(\alpha_i) - \alpha_i\psi(\alpha_{i+1})}{\alpha_{i+1} - \alpha_i} .$$

The derivative for $\phi(\alpha)$ in this interval is

$$\frac{\partial}{\partial\alpha}\,\phi(\alpha) = 1 - \frac{\partial}{\partial\alpha}\,(\psi(\alpha))[\psi(\alpha)]^{-2} .$$

The roots of the derivative are the solution to the equation:

$$[\psi(\alpha)]^2 = \frac{\partial}{\partial\alpha}\,(\psi(\alpha))$$

or

$$\left[\frac{\psi(\alpha_{i+1}) - \psi(\alpha_i)}{\alpha_{i+1} - \alpha_i}\,\alpha + \frac{\alpha_{i+1}\psi(\alpha_i) - \alpha_i\psi(\alpha_{i+1})}{\alpha_{i+1} - \alpha_i}\right]^2 = \frac{\psi(\alpha_{i+1}) - \psi(\alpha_i)}{\alpha_{i+1} - \alpha_i}$$

$$\Rightarrow \frac{\psi(\alpha_{i+1}) - \psi(\alpha_i)}{\alpha_{i+1} - \alpha_i}\,\alpha + \frac{\alpha_{i+1}\psi(\alpha_i) - \alpha_i\psi(\alpha_{i+1})}{\alpha_{i+1} - \alpha_i} = \pm\sqrt{\frac{\psi(\alpha_{i+1}) - \psi(\alpha_i)}{\alpha_{i+1} - \alpha_i}}$$

which yields the two roots

$$\alpha^* = \frac{\alpha_i\psi(\alpha_{i+1}) - \alpha_{i+1}\psi(\alpha_i)}{\alpha_{i+1} - \alpha_i} + \sqrt{\frac{\alpha_{i+1} - \alpha_i}{\psi(\alpha_{i+1}) - \psi(\alpha_i)}}$$

$$\alpha^{**} = \frac{\alpha_i\psi(\alpha_{i+1}) - \alpha_{i+1}\psi(\alpha_i)}{\alpha_{i+1} - \alpha_i} - \sqrt{\frac{\alpha_{i+1} - \alpha_i}{\psi(\alpha_{i+1}) - \psi(\alpha_i)}} .$$

AD-A059 859 CALIFORNIA UNIV BERKELEY OPERATIONS RESEARCH CENTER F/G 12/1
LINEAR DECISION MODELS UNDER RISK.(U)
APR 78 T R HARMS N00014-76-C-0134
UNCLASSIFIED ORC-78-7 NL

2 OF 2
AD
A059859

END
DATE
FILMED
12-78
DDC

(i) If $\psi(\alpha_{i+1}) - \psi(\alpha_i) < 0$, no root exists for the derivative and $\phi(\alpha)$ is monotone increasing for $\alpha \in [\alpha_i, \alpha_{i+1}]$ since

$$\phi(\alpha_i) < \phi(\alpha_{i+1})$$

and

$$\phi(\alpha_i) = \min_{\alpha\varepsilon[\alpha_i,\alpha_{i+1}]} \phi(\alpha) .$$

If $\psi(\alpha_{i+1}) = \psi(\alpha_i)$ then

$$\psi(\alpha) = \psi(\alpha_i) = \psi(\alpha_{i+1}) \quad \text{for} \quad \alpha \in [\alpha_i, \alpha_{i+1}]$$

and $\phi(\alpha) = \alpha + \frac{1}{\psi(\alpha_i)}$ is an increasing linear function and

$$\phi(\alpha_i) = \min_{\alpha\varepsilon[\alpha_i,\alpha_{i+1}]} \phi(\alpha) .$$

(ii) In (a) and (b), α^* and α^{**} are local minimums and as $\phi(\alpha)$ is a convex function for $\alpha \in [\alpha_i, \alpha_{i+1}]$ we have:

For (a): $\phi(\alpha^*) = \min_{\alpha\varepsilon[\alpha_i,\alpha_{i+1}]} \phi(\alpha)$.

For (b): $\phi(\alpha^{**}) = \min_{\alpha\varepsilon[\alpha_i,\alpha_{i+1}]} \phi(\alpha)$.

Note that both α^* and α^{**} cannot belong to the interval at the same time as $\phi(\alpha)$ is convex and does not allow two local minimums.

(c) If neither α^* or $\alpha^{**} \varepsilon [\alpha_i, \alpha_{i+1}]$ then $\phi(\alpha)$ is monotone increasing in the interval since $\phi(\alpha)$ is convex and $\phi(\alpha_i) < \phi(\alpha_{i+1})$. Therefore

$$\phi(\alpha_i) = \min_{\alpha\varepsilon[\alpha_i, \alpha_{i+1}]} \phi(\alpha) . \blacksquare$$

Theorem 5.17:

For an interval $[\alpha_{i-1}, \alpha_i]$ such that $\phi(\alpha_i) < \phi(\alpha_{i-1})$

(i) If $\psi(\alpha_i) - \psi(\alpha_{i-1}) < 0$ then

$$\phi(\alpha_i) = \min_{\alpha\varepsilon[\alpha_{i-1}, \alpha_i]} \phi(\alpha) .$$

(ii) If $\psi(\alpha_i) - \psi(\alpha_{i-1}) > 0$ then one of the following is true:

(a) $\alpha^* \varepsilon [\alpha_{i-1}, \alpha_i]$ and $\phi(\alpha^*) = \min_{\alpha\varepsilon[\alpha_{i-1}, \alpha_i]} \phi(\alpha)$

(b) $\alpha^{**} \varepsilon [\alpha_{i-1}, \alpha_i]$ and $\phi(\alpha^{**}) = \min_{\alpha\varepsilon[\alpha_{i-1}, \alpha_i]} \phi(\alpha)$

(c) $\alpha^* \notin [\alpha_{i-1}, \alpha_i]$ and $\alpha^{**} \notin [\alpha_{i-1}, \alpha_i]$ and

$\phi(\alpha_i) = \min_{\alpha\varepsilon[\alpha_{i-1}, \alpha_i]} \phi(\alpha)$

where:

$$\alpha^* = \frac{\alpha_{i-1}\psi(\alpha_i) - \alpha_i\psi(\alpha_{i-1})}{\alpha_i - \alpha_{i-1}} + \sqrt{\frac{\alpha_i - \alpha_{i-1}}{\psi(\alpha_i) - \psi(\alpha_{i-1})}}$$

$$\alpha^{**} = \frac{\alpha_{i-1}\psi(\alpha_i) - \alpha_i\psi(\alpha_{i-1})}{\alpha_i - \alpha_{i-1}} - \sqrt{\frac{\alpha_i - \alpha_{i-1}}{\psi(\alpha_i) - \psi(\alpha_{i-1})}} .$$

Proof:

In a similar fashion to the proof of Theorem 5.16, it is easy to show that the roots of the derivatives of $\phi(\alpha)$ for the interval $[\alpha_{i-1},\alpha_i]$ are the α^* and α^{**} mentioned in the Theorem 5.16.

(i) If $\psi(\alpha_i) - \psi(\alpha_{i-1}) < 0$ no root of the derivative exists and $\phi(\alpha)$ is monotone decreasing in the interval since $\phi(\alpha)$ is convex and $\phi(\alpha_i) < \phi(\alpha_{i-1})$ so

$$\phi(\alpha_i) = \min_{\alpha\varepsilon[\alpha_{i-1},\alpha_i]} \phi(\alpha) .$$

(ii) The same reasoning as for (ii) of Theorem 5.16 applies here. Note, however, that the case $\psi(\alpha_i) = \psi(\alpha_{i+1})$ does not arise, since then

$$\phi(\alpha_i) = \alpha_i + \frac{1}{\psi(\alpha_i)}$$

and

$$\phi(\alpha_{i-1}) = \alpha_{i-1} + \frac{1}{\psi(\alpha_i)}$$

which contradicts $\alpha_{i-1} < \alpha_i$ and $\phi(\alpha_{i-1}) > \phi(\alpha_i)$.■

G. The Algorithm for Finding the Optimal α

Let us restate the original problem:

$$\text{minimize} \quad c^T x + \frac{1}{d_0 + d^T y}$$

$$\begin{aligned} \text{subject to} \quad Ax - By &\geq b \\ x &\geq 0 \\ 1 \geq y &\geq 0 \end{aligned}$$

with $d \geq 0$ and $d_0 \geq 0$ and the condition that $\exists(\bar{x},\bar{y})$ feasible such that $d_0 + d^T\bar{y} > 0$.

Step 1:

Find a starting point.

1.0:

Solve the initial problem:

$$\text{minimize} \quad c^T x - d^T y - d_0$$

$$\begin{aligned} \text{subject to} \quad Ax - By &\geq b \\ x &\geq 0 \\ 1 \geq y &\geq 0 \; . \end{aligned}$$

This is a linear program and can be solved using the usual algorithm.

1.1:

If the problem is infeasible STOP. The original problem is infeasible (Theorem 5.12).

1.2:

If this problem is unbounded STOP. The original problem is unbounded (Theorem 5.13).

1.3:

Otherwise call the optimal solution $(\bar{x},\bar{y})$ set $c^T\bar{x} = \alpha_o$.
GO TO STEP 2.

Step 2:

2.0:

Consider the problem:

$$\begin{aligned} \psi(\alpha) = \text{-minimize} \quad & -d^T y - d_0 \\ \text{subject to} \quad & Ax - By \geq b \\ & x \geq 0 \\ & 1 \geq y \geq 0 \\ & c^T x \leq \alpha \end{aligned}$$

and $\phi(\alpha) = \alpha + \dfrac{1}{\psi(\alpha)}$.

$\psi(\alpha_o)$ has for optimal solution $(\bar{x},\bar{y})$ the optimal solution of the initial problem (Corollary 5.14).

$\psi(\alpha_o)$ has to be presented in a form ready for parametric analysis (Note 1 at the end of this section shows how the transition from Step 1 to Step 2 can be done easily).

2.0.0:

If $\psi(\alpha_o) = \sum_{j=0}^{n} d_j$, $\psi(\alpha)$ has reached its upper bound and by Corollary 5.10; the function

$$\phi(\alpha) = \alpha + \frac{1}{\sum_{j=0}^{n} d_j}$$

is increasing linearly, so the search has to be done for $\alpha \leq \alpha_o$; hence set $i := 0$. GO TO 2.2.

2.0.1:

If $\psi(\alpha_o) > 0$, set $i := 0$. GO TO 2.1.

2.0.2:

If $\psi(\alpha_o) = 0$, $\phi(\alpha_o)$ is not defined, we need to increase α until $\psi(\alpha) > 0$ is found, such an α exists by Condition 1. Set $i := 0$.

2.0.3:

Find α_{i+1} the next point where the basis changes in the parametric linear program $\psi(\alpha)$ by increasing α from α_i .

2.0.4:

If $\psi(\alpha_{i+1}) = 0$, set $i := i+1$. GO TO 2.0.3.

2.0.5:

If $\psi(\alpha_{i+1}) > 0$, set $\phi(\alpha_i) = +\infty$. Set $i := i+1$. GO TO 2.1.

The algorithm resumes with a new feasible point for which $\phi(\alpha)$ is defined.

<u>2.1</u>:

Find α_{i+1} the next point where the basis changes in the parametric linear program $\psi(\alpha)$ by increasing α from α_i .

<u>2.1.0</u>:

If $\psi(\alpha_{i+1}) = \sum_{j=0}^{n} d_j$, an upper bound for $\psi(\alpha)$ has been reached (Corollary 5.10) and $\phi(\alpha)$ is increasing linearly beyond α_{i+1} .

(a) If $\phi(\alpha_i) > \phi(\alpha_{i+1})$, the optimal interval is $[\alpha_i, \alpha_{i+1}]$. GO TO STEP 3.2 (optimality criterion 1 (i)).

(b) If $\phi(\alpha_i) < \phi(\alpha_{i+1})$, GO TO STEP 2.1.2, the optimality criterion 1 is not satisfied.

<u>2.1.1</u>:

If $\phi(\alpha_{i+1}) < \phi(\alpha_i)$, set $i := i + 1$. GO TO 2.1. Optimality criterion 1 is not satisfied.

<u>2.1.2</u>:

If $\phi(\alpha_{i+1}) > \phi(\alpha_i)$:

<u>*Case 1:*</u> $i = 0$

GO TO 2.2. Only two points have been investigated; the other side of α_o needs to be investigated.

Case 2: $i \neq 0$

The optimal solution is in the interval $[\alpha_{i-1},\alpha_{i+1}]$. GO TO STEP 3. ((i) of optimality criterion 1).

2.1.3:

If $\phi(\alpha_{i+1}) = \phi(\alpha_i)$ STOP. $\alpha \in [\alpha_i,\alpha_{i+1}]$ is an optimal solution ((ii) of optimality criterion 1).

2.2:

Find α_{i-1} the next point for which the basis changes in the parametric linear program $\psi(\alpha)$ when α is decreased from α_i .

If $\psi(\alpha)$ becomes infeasible for $\alpha < \alpha_i$, the optimal solution is in the range $[\alpha_i,\alpha_{i+1}]$. GO TO STEP 3.2. ((iii) of optimality criterion 1).

2.2.0:

If $\psi(\alpha_{i-1}) = 0$, $\psi(\alpha_{i-1})$ has reached its lower bound (Proposition 5.8). Set $\phi(\alpha_{i-1}) = +\infty$.

2.2.1:

If $\phi(\alpha_{i-1}) < \phi(\alpha_i)$, set $i := i - 1$. GO TO 2.2.

2.2.2:

If $\phi(\alpha_{i-1}) > \phi(\alpha_i)$, the optimal solution is in the interval $[\alpha_{i-1},\alpha_i]$. GO TO STEP 3. ((i) of optimality criterion 1).

2.2.3:

If $\phi(\alpha_{i-1}) = \phi(\alpha_i)$ STOP. $\forall \alpha \in [\alpha_{i-1},\alpha_i]$ α is an optimal solution. ((ii) of optimality criterion 1).

Step 3:

Finding the optimal α once an interval has been identified.

3.1:

Interval $[\alpha_{i-1},\alpha_i]$.

3.1.1:

If $\psi(\alpha_i) - \psi(\alpha_{i-1}) < 0$ the solution is either α_i or in the range $[\alpha_i,\alpha_{i+1}]$. GO TO 3.2. (Theorem 5.17 (i)).

3.1.2:

If $\psi(\alpha_i) - \psi(\alpha_{i-1}) > 0$

(a) Compute $\alpha^* = \dfrac{\alpha_{i-1}\psi(\alpha_i) - \alpha_i\psi(\alpha_{i-1})}{\alpha_i - \alpha_{i-1}} + \sqrt{\dfrac{\alpha_i - \alpha_{i-1}}{\psi(\alpha_i) - \psi(\alpha_{i-1})}}$.

If $\alpha^* \in [\alpha_{i-1},\alpha_i]$ STOP. The optimal solution is α^* (Theorem 5.17 (ii)a).

(b) Compute $\alpha^{**} = \dfrac{\alpha_{i-1}\psi(\alpha_i) - \alpha_i\psi(\alpha_{i-1})}{\alpha_i - \alpha_{i-1}} - \sqrt{\dfrac{\alpha_i - \alpha_{i-1}}{\psi(\alpha_i) - \psi(\alpha_{i-1})}}$.

If $\alpha^{**} \in [\alpha_{i-1},\alpha_i]$ STOP. The optimal solution is α^{**} (Theorem 5.17 (ii)b).

(c) The optimal solution is either α_i or in the interval $[\alpha_i, \alpha_{i+1}]$. GO TO 3.2 (Theorem 5.17 (ii)c).

<u>3.2</u>:

Interval $[\alpha_i, \alpha_{i+1}]$.

<u>3.2.1</u>:

If $\psi(\alpha_{i+1}) - \psi(\alpha_i) \leq 0$ STOP. The optimal solution is α_i , by Theorem 5.16 (i) since the other interval has already been investigated.

<u>3.2.2</u>:

If $\psi(\alpha_{i+1}) - \psi(\alpha_i) > 0$

(a) Compute $\alpha^* = \dfrac{\alpha_i \psi(\alpha_{i+1}) - \alpha_{i+1}\psi(\alpha_i)}{\alpha_{i+1} - \alpha_i} + \sqrt{\dfrac{\alpha_{i+1} - \alpha_i}{\psi(\alpha_{i+1}) - \psi(\alpha_i)}}$.

If $\alpha^* \in [\alpha_i, \alpha_{i+1}]$ STOP. The optimal solution is α^* (Theorem 5.16 (ii)a).

(b) Compute $\alpha^{**} = \dfrac{\alpha_i \psi(\alpha_{i+1}) - \alpha_{i+1}\psi(\alpha_i)}{\alpha_{i+1} - \alpha_i} - \sqrt{\dfrac{\alpha_{i+1} - \alpha_i}{\psi(\alpha_{i+1}) - \psi(\alpha_i)}}$.

If $\alpha^{**} \in [\alpha_i, \alpha_{i+1}]$ STOP. The optimal solution is α^{**} (Theorem 5.16 (ii)b).

(c) STOP. The optimal solution is α_i by Theorem 5.16 (ii)c since the other interval has been investigated.

Note 1:

The transition from Step 1 to Step 2 in the previous algorithm can be easily made by doing the following:

1. Modify the initial problem by adding the constraint $c^T x \leq m$ where m is chosen large enough as not to affect the outcome of the problem, (i.e., if at optimality $c^T \bar{x} = m$, then the initial problem is unbounded).

Initial Problem:

$$\begin{aligned} \text{minimize} \quad & c^T x - d^T y - d_0 \\ \text{subject to} \quad & Ax - By \geq b \\ & y \leq 1 \\ & c^T x \leq m \\ & x \geq 0 \quad y \geq 0 . \end{aligned}$$

We will denote this linear program in standard form as:

$$\begin{aligned} \text{minimize} \quad & u^T z \\ \text{subject to} \quad & Mz = t \\ & z \geq 0 . \end{aligned}$$

At optimality the current tableau is denoted:

$$\begin{aligned} \text{minimize} \quad & \bar{u}^T z \\ \text{subject to} \quad & \bar{M} z = \bar{t} \\ & z \geq 0 \end{aligned}$$

with optimal solution $\bar{z} = (\bar{x},\bar{y},\bar{w})$.

2. For Step 2, modify this optimal tableau in the following way:

a) Add $\bar{u}'^T$ to the current objective function $\bar{u}^T$ where:

$$\bar{u}'^T = u' - (u'_s)^T \bar{M}$$

where s is the optimal basis and

$$u' = (-c^T, 0) .$$

b) Add $\bar{t}'$ to the current right-hand side $\bar{t}$ where:

$$\bar{t}' = M_s^{-1} t' \quad \text{and} \quad t' = \begin{bmatrix} 0 \\ \vdots \\ -m + \alpha_o \end{bmatrix} .$$

M_s^{-1} is the basic inverse with respect to the basis s .

$-m + \alpha_o$ in t' corresponds to the constraint $c^T x \leq m$.

3. The new tableau thus obtained should be either optimal and ready for parametric analysis or a few more iterations might be needed to achieve that state.

H. Finding the Optimal Solution to the Original Problem from the Optimal α

Once an optimal $\bar{\alpha}$ has been identified, Theorem 5.1 applied to our problem, tells us that an optimal $(\bar{x},\bar{y})$ can be found by solving the problem.

$$\text{minimize} \quad \frac{1}{d_0 + d^T y}$$

$$\begin{aligned} \text{subject to} \quad Ax - By &\geq b \\ x &\geq 0 \\ 1 \geq y &\geq 0 \\ c^T x &\leq \bar{\alpha} \ . \end{aligned}$$

However, Proposition 5.6 tells us that $(\bar{x},\bar{y})$ is also the optimal solution of the problem:

$$\psi(\bar{\alpha}) = -\text{minimize} \quad -d_0 - d^T y$$

$$\begin{aligned} \text{subject to} \quad Ax - By &\geq b \\ x &\geq 0 \\ 1 \geq y &\geq 0 \\ c^T x &\leq \bar{\alpha} \ . \end{aligned}$$

Since we have information about the problem $\psi(\alpha)$ we can use it to find $(\bar{x},\bar{y})$.

1. If $\bar{\alpha} = \alpha_k$, $\alpha_k \in \{\alpha_i\}$, the sequence generated in the algorithm, then $(\bar{x},\bar{y}) = (x^k,y^k)$. The optimal solution of $\psi(\alpha_k)$ is readily available.
2. If $\alpha_k < \bar{\alpha} < \alpha_{k+1}$ and $\alpha_k \in \{\alpha_i\}$, $\alpha_{k+1} \in \{\alpha_i\}$, then:

$$(\bar{x},\bar{y})_{s^k} = (x^k,y^k)_{s^k} + (\bar{\alpha} - \alpha_k)M_{s^k}^{-1}\begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}$$

and $(\bar{x},\bar{y})_G = 0$.

s_k is the optimal basic sequence for $\psi(\alpha_k)$; (x^k,y^k) is the optimal solution of $\psi(\alpha_k)$; $M^{-1}_{s^k}$ is the basic inverse at optimality of $\psi(\alpha_k)$ and G is the set of indices of the non-basic variables. The one in the vector $\begin{bmatrix} 0 \\ \vdots \\ 1 \end{bmatrix}$ corresponds to the constraint $c^T x \leq \alpha$.

This is just an application of parametric linear programming where α is the parameter.

I. Convergence Property of the Algorithm

We will need the following propositions to establish the convergence of the algorithm.

Proposition 5.18:

If the linear program $\psi(\alpha)$ is nondegenerate, then the sequence $\{\alpha_i\}$ generated in the algorithm is finite.

Proof:

$\{\alpha_i\}$ is the sequence of real numbers for which the basis changes in the parametric linear program $\psi(\alpha)$. There is a one to one correspondence between the sequence $\{\alpha_i\}$ and the sequence of basic sequences of the problem $\psi(\alpha)$. This sequence (of basic sequences) is finite since there is a finite number of such sequences and none is repeated, because under nondegeneracy the objective function is improved at each iteration. Therefore, the sequence $\{\alpha_i\}$ is finite.■

Proposition 5.19:

Under degeneracy of $\psi(\alpha)$, the sequence $\{\alpha_i\}$ is finite.

Proof:

In this case, a lexicographical scheme or another appropriate rule of pivoting has to be used in solving the parametric linear program $\psi(\alpha)$ to insure that none of the basic sequences is repeated as is mentioned in the proof of Proposition 5.18.■

We can now prove the finiteness of the algorithm.

Theorem 5.20:

The algorithm terminates in a finite number of steps.

Proof:

The initial problem terminates in a finite number of steps since it is a linear program and the simplex method terminates in a finite number of steps (degeneracy is taking care of with a lexicographical scheme). The algorithm then proceeds to generate the sequence $\{\alpha_i\}$ from a second linear program and Propositions 5.18 and 5.19 tell us that this sequence is finite. Once an interval is identified from this sequence in Phase 2, a small finite number of steps is necessary to obtain the optimal solution.■

We have just seen that the algorithm terminates in a finite number of steps, however how many steps are needed to obtain the optimal solution is also of interest. An upper bound to this number of steps could be calculated, but as in the simplex method, would not be indicative of the real amount of work needed for the algorithm. The best way to get a feel for this amount is to compare it to the simplex method. Essentially, the initial problem represents one whole linear programming problem, then it is modified, and parametric analysis is performed. We can confidently say that the amount of work is of the order of the solving one linear program and performing parametric analysis on the right-hand side.

CHAPTER VI

In this chapter the models I and II are generalized to become joint-chance constrained programming problems. Some approximations are presented using the concept of associativity of random variables and the results of Chapter II.

A. Generalized Models I and II

We would like to set more general constraints on the feasibility sets of Models I and II. For example, we could add the constraint that the probability of the solution being feasible be greater than a certain specified number. This is why we introduced joint probability constraints; joint probability constraints are found in the literature in [1] and [10]; we will use these in the following generalized models I and II.

Generalized Model I:

$$\text{minimize} \quad E\left\{\sum_{j=1}^{n} c_j(x_j + e_j)\right\}$$

$$\text{subject to} \quad x \in \Omega$$

$$\text{Prob}\left\{\sum_{j=1}^{n} A_{ij}(x_j + e_j) \geq b_i\right\} \geq \gamma_i \qquad i = 1, \ldots, m$$

$$\text{Prob}\left\{\bigcap_{i \in s_k}\left\{\sum_{j=1}^{n} A_{ij}(x_j + e_j) \geq b_i\right\}\right\} \geq \gamma_k \qquad k = m+1, \ldots, t$$

where s is the set of indices of the rows of A $s = \{1, \ldots, m\}$, and s_k are subsets of s for $k = m+1, \ldots, t$.

Ω is a polyhedral set defined by a set of linear constraints.

Generalized Model II:

$$\text{minimize} \quad E\left\{\sum_{j=1}^{n} c_j(1+\alpha_j)x_j\right\}$$

$$\text{subject to} \quad x \in \Omega$$

$$\text{Prob}\left\{\sum_{j=1}^{n} A_{ij}(1+\alpha_j)x_j \geq b_i\right\} \geq \gamma_i \qquad i = 1, \ldots, m$$

$$\text{Prob}\left\{\bigcap_{i \in s_k}\left\{\sum_{j=1}^{n} A_{ij}(1+\alpha_j)x_j \geq b_i\right\}\right\} \geq \gamma_k \qquad k = m+1, \ldots, t\ .$$

s , s_k and Ω are defined as in Model I.

The generalized models I and II fall in the category of chance-constrained programming with joint constraints. However, the constraints are not statistically independent since the e_j and α_j are involved in all the constraints. Therefore, these problems cannot be treated as described in [1] and [10]. We will draw upon the theory of associated random variables developed by Esary-Proschan and Walkup [6] to determine conservative approximations to these problems.

B. Associated Random Variables

Definition:

Random variables X_j $j = 1, \ldots, n$ are associated if

$$\text{Cov}\,(\Gamma(X), \Delta(X)) \geq 0$$

for all pairs of increasing binary functions Γ , Δ where $X = (X_1, \ldots, X_n)$.

Properties of associated random variables:

Theorem 6.1:

Increasing functions of associated random variables are associated.

Proof:

See Barlow's book on Reliability, Chapter 2 [2].■

Theorem 6.2:

Independent random variables are associated.

Proof:

See Barlow's book on Reliability, Chapter 2 [2].

Theorem 6.3:

If $X_1, X_2, \ldots, X_n$ are associated random variables, then for all $(x_1, \ldots, x_n)$:

$$\text{Prob}\,[X_1 > x_1, \ldots, X_n > x_n] \geq \prod_{i=1}^{n} \text{Prob}\,[X_i > x_i]$$

$$\text{Prob}\,[X_1 \leq x_1, \ldots, X_n \leq x_n] \geq \prod_{i=1}^{n} \text{Prob}\,[X_i \leq x_i] \, .$$

Proof:

See Barlow's book on Reliability, Chapter 2 [2].■

We will now use the previous three theorems to prove theorems relevant to our models. We will first define the following set of matrices.

Definition:

Let $\mathcal{M}$ be the set of matrices for which all the elements in any column have the same sign.

We can now present the following results.

Theorem 6.4:

If X_j $j = 1, \ldots, n$ are independent random variables and A is an $m \times n$ matrix belonging to $\mathcal{M}$, then the random variables $\sum_{j=1}^{n} A_{ij} X_j$ $i = 1, \ldots, m$ are associated.

Proof:

Call T the set of indices of the column of $A : T = \{1, \ldots, n\}$ and s the set of indices of the rows of $A : s = \{1, \ldots, m\}$.

Consider the following sets

$$T' = \{j \mid A_{ij} \leq 0 \text{ for all } i \in s\}$$

$$T'' = \{j \mid A_{ij} \geq 0 \text{ for all } i \in s\} .$$

As $A \in \mathcal{M} : T' \cup T'' = T$.

Now consider the set of independent random variables U comprised of X_j for $j \in T'$ and $-X_j$ for $j \in T''$.

By Theorems 6.1 and 6.2 the random variables $\sum_{j\varepsilon T'} |A_{ij}|(-X_j) + \sum_{j\varepsilon T''} |A_{ij}|X_j$ $i = 1, \ldots, m$ are associated since U is a set of independent random variables therefore associated by Theorem 6.2, and $\sum_{j=1}^{n} |A_{ij}|X_j$ is an increasing function of X_j which makes these random variables associated by Theorem 6.1. However, for $i = 1, \ldots, m$, $\sum_{j\varepsilon T'} |A_{ij}|(-X_j) + \sum_{j\varepsilon T''} |A_{ij}|X_j = \sum_{j=1}^{n} A_{ij}X_j$ by definition of T' and T''. ■

Theorem 6.5:

If X_j $j = 1, \ldots, n$ are associated random variables and A is a nonnegative $m \times n$ matrix then the random variables $\sum_{j=1}^{n} A_{ij}X_j$ $i = 1, \ldots, m$ are associated.

Proof:

As $A \geq 0$, for $i = 1, \ldots, m$, $\sum_{j=1}^{n} A_{ij}x_j$ are increasing functions of x; then by Theorem 6.1 for $i = 1, \ldots, m$, $\sum_{j=1}^{n} A_{ij}X_j$ are associated random variables. ■

C. Conservative Approximations for the Generalized Models I and II

Using the results of Section B, we shall first present the following approximations.

Theorem 6.6:

In Model I, if either (i) or (ii) below holds

(i) e_j $j = 1, \ldots, n$ are associated random variables and $A_{s_k} \geq 0$ for $k = m+1, \ldots, t$ (where A_{s_k} is the matrix consisting of the rows i of A such that $i \in s_k$).

(ii) e_j $j = 1, \ldots, n$ are independent random variables and $A_{s_k} \in M$ for $k = m+1, \ldots, t$.

Then, if x satisfies:

$$\prod_{i \in s_k} \left(1 - \text{Prob}\left[\sum_{j=1}^{n} A_{ij} e_j \leq b_i - \sum_{j=1}^{n} A_{ij} x_j\right]\right) \geq \gamma_k \qquad k = m+1, \ldots, t$$

then x satisfies:

$$\text{Prob}\left\{\bigcap_{i \in s_k} \left\{\sum_{j=1}^{n} A_{ij}(x_j + e_j) \geq b_i\right\}\right\} \geq \gamma_k \qquad k = m+1, \ldots, t \ .$$

<u>Proof</u>:

(i) Theorem 6.5 tells us that for $i \in s_k$ $\sum_{j=1}^{n} A_{ij} e_j$ are associated random variables. Applying Theorem 6.3 we have:

$$\text{Prob}\left\{\bigcap_{i \in s_k} \left\{\sum_{j=1}^{n} A_{ij} e_j \geq z\right\}\right\} \geq \prod_{i \in s_k} \text{Prob}\left\{\sum_{j=1}^{n} A_{ij} e_j \geq z\right\} .$$

Hence if x is such that:

$$\prod_{i \in s_k} \left(1 - \text{Prob}\left\{\sum_{j=1}^{n} A_{ij} e_j \leq b_i - \sum_{j=1}^{n} A_{ij} x_j\right\}\right) \geq \gamma_k$$

then

$$\text{Prob}\left\{\bigcap_{i\varepsilon s_k}\left\{\sum_{j=1}^{n} A_{ij}x_j + \sum_{j=1}^{n} A_{ij}e_j \geq b_i\right\}\right\} \geq \gamma_k .$$

(ii) Theorem 6.4 tells us that for $i \varepsilon s_k$, $\sum_{j=1}^{n} A_{ij}e_j$ are associated random variables, from now on the reasoning is the same as in (i).■

Theorem 6.7:

In Model II if either (i) or (ii) below holds:

(i) α_j $j = 1, \ldots, n$ are associated random variables and $A_{s_k} \geq 0$ for $k = m + 1, \ldots, t$.

(ii) α_j $j = 1, \ldots, n$ are independent random variables and $A_{s_k} \varepsilon \mathcal{M}$ for $k = m + 1, \ldots, t$.

Then if x satisfies:

$$\prod_{i\varepsilon s_k}\left(1 - \text{Prob}\left\{\sum_{j=1}^{n} A_{ij}x_j\alpha_j \leq b_i - \sum_{j=1}^{n} A_{ij}x_j\right\}\right) \geq \gamma_k \qquad k = m + 1, \ldots, t$$

x satisfies

$$\text{Prob}\left\{\bigcap_{i\varepsilon s_k}\left\{\sum_{j=1}^{n} A_{ij}(1 + \alpha_j)x_j \geq b_i\right\}\right\} \geq \gamma_k \qquad k = m + 1, \ldots, t .$$

Proof:

(i) As $A_{s_k} \geq 0$, the matrix $\bar{A}_{s_k} \in M$ where $\left(\bar{A}_{s_k}\right)_{ij} = \left(A_{s_k}\right)_{ij} x_j$ for all i and j . Therefore, by Theorem 6.4 $i \in s_k$, $\sum_{j=1}^{n} A_{ij}x_j\alpha_j$ are associated random variables.

Similarly under (ii), as $A_{s_k} \in M$, $\bar{\bar{A}}_{s_k} \in M$ where $\left(\bar{\bar{A}}_{s_k}\right)_{ij} = \left(A_{s_k}\right)_{ij} x_j$ and by Theorem 6.4 for $i \in s_k$ $\sum_{j=1}^{n} A_{ij}x_j\alpha_j$ are associated random variables.

Now following the proof of Theorem 6.6, by Theorem 6.3 we have

$$\text{Prob}\left\{\bigcap_{i\in s_k}\left\{\sum_{j=1}^{n} A_{ij}x_j\alpha_j \geq z\right\}\right\} \geq \prod_{i\in s_k} \text{Prob}\left\{\sum_{j=1}^{n} A_{ij}x_j\alpha_j \geq z\right\} .$$

Hence if x is such that:

$$\prod_{i\in s_k}\left(1 - \text{Prob}\left\{\sum_{j=1}^{n} A_{ij}x_j\alpha_j \leq b_i - \sum_{j=1}^{n} A_{ij}x_j\right\}\right) \geq \gamma_k$$

then:

$$\text{Prob}\left\{\bigcap_{i\in s_k}\left\{\sum_{j=1}^{n} A_{ij}(1 + \alpha_j)x_j \geq b_i\right\}\right\} \geq \gamma_k . \blacksquare$$

We shall now use Theorem 6.6 and 6.7 to present complete formulations of the conservative approximations for Models I and II.

1. Model I:

Theorem 6.8:

For either condition (i) or (ii) of Theorem 6.6 the following problem is a conservative approximation to the generalized Model I.

$$\text{minimize} \quad c^T x + \sum_{j=1}^{n} c_j E\{e_j\}$$

$$\text{subject to} \quad x \in \Omega$$

$$\sum_{j=1}^{n} A_{ij} x_j \geq b_i - G_i^{-1}(1 - \gamma_i) \qquad i = 1, \ldots, n$$

$$\prod_{i \varepsilon s_k} \left(1 - G_i \left[b_i - \sum_{j=1}^{n} A_{ij} x_j \right] \right) \geq \gamma_k \qquad k = m + 1, \ldots, t$$

where $G_i(z) = \text{Prob}\left\{ \sum_{j=1}^{n} A_{ij} e_j \leq z \right\}$.

Proof:

This model is valid by Theorem 6.6 and the deterministic equivalent of Model I presented in Chapter I.■

2. Model II:

Theorem 6.9:

For either condition (i) or (ii) of Theorem 6.7, the following problem is a conservative approximation to the generalized Model II.

$$\text{minimize} \quad \sum_{j=1}^{n} c_j(1 + E(\alpha_j))x_j$$

$$\text{subject to} \quad x \in \Omega$$

$$\sum_{j=1}^{n} A_{ij}x_j + {}_iG_x^{-1}(1 - \gamma_i) \geq b_i \qquad i = 1, \ldots, m$$

$$\prod_{i \in s_k} \left(1 - {}_iG_x\left[b_i - \sum_{j=1}^{n} A_{ij}x_j \right] \right) \geq \gamma_k \qquad k = m + 1, \ldots, t$$

where $${}_iG_x(z) = \text{Prob}\left\{ \sum_{j=1}^{n} A_{ij}x_j\alpha_j \leq z \right\} .$$

<u>Proof</u>:

This is the deterministic equivalent presented in Chapter I with the approximation of Theorem 6.7 applied to it.■

D. Conservative Approximations for the Generalized Models I and II Using the Results of Chapter II

If in addition to the conditions outlined in the previous section, we utilize random variables e_j and α_j belonging to the set C defined in Chapter II, explicit conservative approximations can be identified.

1. Model I:

<u>Theorem 6.10</u>:

If for e_j $j = 1, \ldots, n$ $e_j \in C$ either the following (i) or (ii) holds:

(i) e_j $j = 1, \ldots, n$ are associated random variables and

$A_{s_k} \geq 0$ for $k = m + 1, \ldots, t$.

(ii) e_j $j = 1, \ldots, n$ are independent random variables and

$A_{s_k} \in M$ for $k = m + 1, \ldots, t$.

Then for $\gamma \geq \frac{1}{2}$ the following convex program is a conservative approximation to the generalized Model I.

$$\text{minimize} \quad c^T x$$

$$\text{subject to} \quad x \in \Omega$$

$$y_i - \sum_{j=1}^{n} \frac{A_{ij}}{2 \sum_{j=1}^{n} |A_{ij}| a_j} x_j = \frac{1}{2} - \frac{b_i}{2 \sum_{j=1}^{n} |A_{ij}| a_j} \qquad i = 1, \ldots, m$$

$$y_i \geq \gamma_i \qquad i = 1, \ldots, m$$

$$\sum_{i \in s_k} \mathrm{Ln}(y_i) \geq \mathrm{Ln}(\gamma_k) \qquad k = m + 1, \ldots, t$$

where the $[-a_j, a_j]$ $j = 1, \ldots, n$ are defined to be the finite ranges of e_j $j = 1, \ldots, n$.

Proof:

1) As $e_j \in C$ for $j = 1, \ldots, n$ the objective function is simply $c^T x$.

2) As shown for Model I in Theorem 3.1 the first set of constraints result of the inequality for $\gamma \geq \frac{1}{2}$

$$G_i\left[\sum_{j=1}^{n} A_{ij}x_j - b_i\right] \geq \frac{\left(\sum_{j=1}^{n} A_{ij}x_j - b_i\right) + \sum_{j=1}^{n} |A_{ij}|a_j}{2\sum_{j=1}^{n} |A_{ij}|a_j} .$$

If we call the right-hand side y_i then it is obvious that if $y_i \geq \gamma_i$ then

$$G_i\left[\sum_{j=1}^{n} A_{ij}x_j - b_i\right] \geq \gamma_i$$

for all $i = 1, \ldots, m$.

3) Theorem 6.8 yields the first approximation.

$$\text{Prob}\left\{\bigcap_{i\varepsilon s_k}\left[\sum_{j=1}^{n} A_{ij}(x_j + e_j) \geq b_i\right]\right\} \geq \prod_{i\varepsilon s_k} G_i\left[b_i - \sum_{j=1}^{n} A_{ij}x_j\right] .$$

However, since $e_j \varepsilon C$ $j = 1, \ldots, n$ then by Corollary 2.10 $\sum_{j=1}^{n} A_{ij}e_j \varepsilon C$ and therefore it symmetric and

$$1 - G_i\left[b_i - \sum_{j=1}^{n} A_{ij}x_j\right] = G_i\left[\sum_{j=1}^{n} A_{ij}x_j - b_i\right] .$$

Using again the same inequality as in part 2) since $\gamma \geq \frac{1}{2}$ we obtain that

$$\text{Prob}\left\{\bigcap_{i\varepsilon s_k}\left[\sum_{j=1}^{n} A_{ij}(x_j + e_j) \geq b_i\right]\right\} \geq \prod_{i\varepsilon s_k} y_i .$$

Hence for $k = m + 1, \ldots, t$ the following constraints are conservative approximations.

$$\prod_{i \varepsilon s_k} y_i \geq \gamma_k$$

or equivalently

$$\sum_{i \varepsilon s_k} Ln(y_i) \geq Ln(\gamma_k) \quad \text{as} \quad y_i \geq \gamma_i \geq 0 \quad i \varepsilon s_k$$

which makes the whole program convex since $\sum_{i \varepsilon s_k} Ln(y_i)$ is a concave function of y_i .■

2. Model II:

Theorem 6.11:

If for $\alpha_j \varepsilon C$ $j = 1, \ldots, n$ either (i) or (ii) holds and $\{x \mid x \geq 0\} \subset \Omega$.

(i) α_j $j = 1, \ldots, n$ are associated and $A_{s_k} \geq 0$ for $k = m + 1, \ldots, t$.

(ii) α_j $j = 1, \ldots, n$ are independent and $A_{s_k} \varepsilon M$ for $k = m + 1, \ldots, t$.

Then for $\gamma \geq \frac{1}{2}$, the following nonlinear program is a conservative approximation to the generalized Model II.

$$\text{minimize} \quad c^T x$$

$$\text{subject to} \quad x \varepsilon \Omega$$

$$\sum_{j=1}^{n} [A_{ij} + |A_{ij}| a_j] x_j - w_i \left[2 \sum_{j=1}^{n} |A_{ij}| a_j x_j \right] = b_i \qquad i = 1, \ldots, m$$

$$w_i \geq \gamma_i \qquad i = 1, \ldots, m$$

$$\prod_{i \in s_k} w_i \geq \gamma_k \qquad k = m+1, \ldots, t .$$

<u>Proof</u>:

As $\alpha_j \in C$ for $j = 1, \ldots, n$ by Corollary 2.10

$\sum_{j=1}^{n} A_{ij} x_j \alpha_j \in C$

$$\text{Prob}\left[\sum_{j=1}^{n} A_{ij}(1+\alpha_j)x_j \geq b_i\right] \geq \gamma_i$$

$$\Longleftrightarrow 1 - \text{Prob}\left[\sum_{j=1}^{n} A_{ij} x_j \alpha_j \leq b_i - \sum_{j=1}^{n} A_{ij} x_j\right] \geq \gamma_i .$$

As $\sum_{j=1}^{n} A_{ij} x_j \alpha_j \in C$

$$\Longleftrightarrow \text{Prob}\left[\sum_{j=1}^{n} A_{ij} x_j \alpha_j \leq \sum_{j=1}^{n} A_{ij} x_j - b_i\right] \geq \gamma_i .$$

Applying Theorem 2.13

$$\text{Prob}\left[\sum_{j=1}^{n} A_{ij} x_j \alpha_j \leq \sum_{j=1}^{n} A_{ij} x_j - b_i\right] \geq \frac{\sum_{j=1}^{n} A_{ij} x_j - b_i + \sum_{j=1}^{n} |A_{ij}| a_j x_j}{2 \sum_{j=1}^{n} |A_{ij}| a_j x_j}$$

since for $\gamma \geq \frac{1}{2} \Rightarrow \sum_{j=1}^{n} A_{ij} x_j - b_i \geq 0$ and $F_j^{-1}(0) = -|A_{ij}| a_j x_j$ for $F(z) = \text{Prob}\{A_{ij} x_j \alpha_j \leq z\}$ by Theorem 2.7.

Therefore if for $x \geq 0$

$$w_j = \frac{\sum_{j=1}^{n} A_{ij}x_j - b_i + \sum_{j=1}^{n} |A_{ij}|a_j x_j}{2 \sum_{j=1}^{n} |A_{ij}|a_j x_j} \geq \gamma_i \ .$$

Then $\text{Prob}\left\{ \sum_{j=1}^{n} A_{ij}(1 + \alpha_j)x_j \geq b_i \right\} \geq \gamma_i$.

Theorem 6.7 showed that

$$\text{Prob}\left(\bigcap_{i\varepsilon s_k} \left[\sum_{j=1}^{n} A_{ij}(1 + \alpha_j)x_j \geq b_i \right] \right) \geq \prod_{i\varepsilon s_k} \text{Prob}\left\{ \sum_{j=1}^{n} A_{ij}(1+\alpha_j)x_j \right\} \geq b_i \ .$$

Therefore if for $x \geq 0$

$$\prod_{i\varepsilon s_k} w_j \geq \gamma_k \ .$$

Then

$$\prod_{i\varepsilon s_k} \text{Prob}\left[\sum_{j=1}^{n} A_{ij}(1 + \alpha_j)x_j \geq b_i \right] \geq \gamma_k$$

and

$$\text{Prob}\left(\bigcap_{i\varepsilon s_k} \left[\sum_{j=1}^{n} A_{ij}(1 + \alpha_j)x_j \geq b_i \right] \right) \geq \gamma_k \ . \blacksquare$$

BIBLIOGRAPHY

[1] Balintfy, J. L., "Nonlinear Programming for Models with Joint Chance-Constraints," in INTEGER AND NONLINEAR PROGRAMMING, J. Abadie, (ed.), North Holland Publishers, Amsterdam, Netherlands, (1970).

[2] Barlow, R. and F. Proschan, STATISTICAL THEORY OF RELIABILITY AND LIFE TESTING: PROBABILITY MODELS, Holt, Rinehart and Winston, (1976).

[3] Charnes, A and W. W. Cooper, "Chance Constrained Programming," <u>Management Science</u>, Vol. 6, pp. 73-79, (1959).

[4] _______________ and _______________, "Deterministic Equivalents for Optimizing and Satisficing Under Chance Constraints," <u>Operations Research</u>, Vol. 11, pp. 18-39, (1963).

[5] Dantzig, G. B., LINEAR PROGRAMMING AND EXTENSIONS, Princeton University Press, Princeton, New Jersey, (1963).

[6] Esary, J. D., F. Proschan and D. Walkup, "Association of Random Variables, With Applications," <u>Annals of Mathematical Statistics</u>, Vol. 38, pp. 1466-1474, (1967).

[7] Feller, W., AN INTRODUCTION TO PROBABILITY THEORY AND ITS APPLICATIONS, VOL. II, Wiley, New York, (1966).

[8] Kataoka, S., "A Stochastic Programming Model," <u>Econometrica</u>, Vol. 31, pp. 181-196, (1963).

[9] Lawrence, J., "Parametric Programming With Extensions to Large Scale Algorithms," ORC 73-18, Operations Research Center, University of California, Berkeley, (1973).

[10] Miller, B. L. and H. M. Wagner, "Chance-Constrained Programming With Joint Constraints," Operations Research, Vol. 13, pp. 930-945, (1965).

[11] Sengupta, J. K., STOCHASTIC PROGRAMMING: METHODS AND APPLICATIONS, North Holland, Amsterdam, (1972),

[12] Stancu-Minason, I. M. and M. J. Wets, "A Research Bibliography in Stochastic Programming 1955-1975," Operations Research, Vol. 24, pp. 1078-1119, (1976).

[13] Vajda, S., PROBABILISTIC PROGRAMMING, Academic Press, New York, (1972).

[14] Walters, D., "Multi-Parametric Mathematical Programming Problems," ORC 76-5, Operations Research Center, University of California, Berkeley, (1976).

APPENDIX NOTATION

The vectors mentioned in the text (i.e., b, c, d) are column vectors; superscript T denotes the transpose of a vector:

$$b = \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix} ; \quad b^T = (b_1, \ldots, b_n) .$$

If A is an $n \times m$ matrix then:

A_{ij} = the element in the i^{th} row, j^{th} column of A

$A_{\cdot j}$ = j^{th} column of A

$A_{i\cdot}$ = i^{th} row of A .

$x \in X$	x is an element of the set X
$\exists x$	there exist at least one x
$X \subset Y$	set X is contained in the set Y
$X \cup Y$	the union of the sets X and Y
$X \cap Y$	intersection of the sets X and Y
$\blacksquare$	end of proof
$\lvert x \rvert$	absolute value of x
$\emptyset$	the empty set
$\{x \mid P\}$	the set of x having property P
$\{x_i\}$	the set $\{\cdots x_{i-1}, x_i, x_{i+1}, \cdots\}$
$\sum_{i=1}^{n} x_i$	$x_1 + x_2 + \cdots + x_n$
$\prod_{i=1}^{n} x_i$	$x_1 x_2 \cdots x_n$

(a,b)	open interval from a to b
$[a,b]$	closed interval from a to b
[n]	reference n in bibliography
/	(through a symbol, e.g., $\neq$) negation
$\Longleftrightarrow$	if and only if
$\Rightarrow$	implies
$\mathrm{Ln}(x)$	natural logarithm of x
$\mathrm{Prob}\{E\}$	probability of the event E occurring
$E\{X\}$	expectation of the random variable X
$f * g$	convolution of the densities f and g
$\mathop{*}\limits_{i=1}^{n} f_i$	$f_1 * f_2 * \cdots * f_n$
F^{-1}	the inverse of the function F
A^{-1}	the inverse of the matrix A .

If γ is a vector the statement

$\gamma = a$ means $\gamma = \begin{bmatrix} a \\ \vdots \\ a \end{bmatrix}$

$\gamma \geq a$ means $\gamma \geq \begin{bmatrix} a \\ \vdots \\ a \end{bmatrix}$.

$I = \begin{bmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{bmatrix}$ the identity matrix

Δf the gradient of the function f .