D1-82-0411

BOEING SCIENTIFIC RESEARCH LABORATORIES

Programming Under Uncertainty: The Equivalent Convex Program

Roger Wets

Mathematics Research

February, 1965

CLEARINGHOUSE FOR FEDERAL SCIENTIFIC AND TECHNICAL INFORMATION, CFSTI
INPUT SECTION 410.11

LIMITATIONS IN REPRODUCTION QUALITY OF TECHNICAL ABSTRACT BULLETIN DOCUMENTS, DEFENSE DOCUMENTATION CENTER (DDC)

AD612896

☐ 1. AVAILABLE ONLY FOR REFERENCE USE AT DDC FIELD SERVICES. COPY IS NOT AVAILABLE FOR PUBLIC SALE.

☒ 2. AVAILABLE COPY WILL NOT PERMIT FULLY LEGIBLE REPRODUCTION. REPRODUCTION WILL BE MADE IF REQUESTED BY USERS OF DDC.

☒ A. COPY IS AVAILABLE FOR PUBLIC SALE.

☐ B. COPY IS NOT AVAILABLE FOR PUBLIC SALE.

☐ 3. LIMITED NUMBER OF COPIES CONTAINING COLOR OTHER THAN BLACK AND WHITE ARE AVAILABLE UNTIL STOCK IS EXHAUSTED. REPRODUCTIONS WILL BE MADE IN BLACK AND WHITE ONLY.

TSL-121-2/64

DATE PROCESSED: 4-2-65

PROCESSOR: GL

D1-82-0411

PROGRAMMING UNDER UNCERTAINTY: THE EQUIVALENT CONVEX PROGRAM

by

Roger Wets

Mathematical Note No. 392

Mathematics Research Laboratory

BOEING SCIENTIFIC RESEARCH LABORATORIES

February 1965

ABSTRACT

This paper is an attempt to describe and characterise the equivalent convex program of a two-stage linear program under uncertainty. The study has been divided into two parts. In the first one, we examine the properties of the solution set of the problem and derive explicit expressions for some particular cases. The second section is devoted to the derivation of the objective function of the equivalent convex program. We show that it is convex and continuous. We also give a necessary condition for its differentiability and establish necessary and sufficient conditions for the solvability of the problem. Finally, we give the equivalent convex program of certain classes of programming under uncertainty problems, i.e. when the constraints and the probability space have particular structures.

BLANK PAGE

I. INTRODUCTION

The <u>standard form</u> of the problem -- which we assume solvable -- to be considered in this paper is:

$$(1)\qquad \begin{array}{lll} \text{Minimize} & z(x) = cx + E_{\xi}\{qy\} & \\ \text{subject to} & Ax \qquad\quad = b & \\ & Tx + \quad My = \xi, & \xi \text{ on } (\Xi, \mathcal{F}, F) \\ & x \geq 0, \quad y \geq 0 & \end{array}$$

where A is a matrix $m \times n$, T is $\bar{m} \times n$, M is $\bar{m} \times \bar{n}$, ξ is a random vector defined on the probability space $(\Xi, \mathcal{F}, F)$.

This problem belongs to the class of stochastic linear programming problems for which one seeks a <u>here-and-now solution</u>. Problem (1) is known in the literature as the <u>two-stage linear program under uncertainty</u>. One interprets it as follows: The decision maker must select the activity levels for x, say $x = \hat{x}$, he then observes the random event $\xi = \hat{\xi}$ and he is finally allowed to take a corrective action y, such that $y \geq 0$, $My = \hat{\xi} - T\hat{x}$ and qy is minimum. This second stage decision y is taken when no uncertainties are left in the problem.

It is clear that we could also write the objective function of (1) as:

$$(1') \qquad z(x) = cx + E_{\xi}\{\text{Min } qy \mid x\}$$

The interpretation given above indicates that (1) as well as (1') are conventional ways to express the same concept. Many practical problems can be formulated to fit the standard form, e.g. inventory problems, planning problems, transportation problems with uncertain demand, etc.

All quantities considered here belong to the reals, denoted $\mathfrak{R}$. Vectors will belong to finite-dimensional spaces $\mathfrak{R}^n$ and whether they are to be regarded as row vectors or column vectors will always be clear from the context in which they appear. Thus, for example, the expressions

$$x = (x_1, x_2, \ldots, x_i, \ldots, x_n)$$

$$Tx = \chi$$

$$y^+ y^- = \sum_{i=1}^{\bar{m}} y_i^+ y_i^-$$

are easily understood. No special provisions have been made for transposing vectors.

For the sake of simplicity, we shall assume that $(\Xi, \mathcal{F}, F)$ is the probability space induced in $\mathfrak{R}^{\bar{m}}$, F determines a Lebesgue-Stieltjes measure and $\mathcal{F}$ is the completion for F of the Borel algebra in $\mathfrak{R}^{\bar{m}}$. We also assume that $\bar{\xi} = E\{\xi\}$ exists. Also, note that our notation ξ on $(\Xi, \mathcal{F}, F)$ is meant to imply that the first stage decision has no effect on the probability space on which ξ is defined. In other words, ξ is independent of x.

The marginal probability space for $i = 1, \ldots, \bar{m}$ will be denoted by $(\Xi_i, \mathcal{F}_i, F_i)$. If it exists, we denote the density function of ξ_i

by $f_1(\xi_1)$. If ξ_1 is a discrete random variable, we denote its probability mass function also by $f_1(\xi_1)$. No confusion should arise from this abuse of notation. Moreover, let α_1 and β_1 be respectiv ly the greatest lower bound and the least upper bound of Ξ_1. If Ξ_1 is not bounded below, we set $\alpha_1 = -\infty$, if Ξ_1 is not bounded above, we set $\beta_1 = +\infty$.

We usually think of Ξ as the convex hull of all elements of $\mathcal{F}$ with positive measure. In other words, Ξ is equal to the intersection of all convex subsets of $\Re^{\bar{m}}$ with measure one. The probability measure may be discrete, continuous, or a mixture of both. Only in one particular case (II.A), shall we use another characterization of Ξ, namely: $\Xi = \{\xi | f(\xi) \neq 0\}$.

The first part of this paper characterizes the solution set of (1), and it points out some of its properties. In the second part, we derive a programming problem whose set of optimal solutions is identical to the set of optimal solutions to problem (1).

II. THE SOLUTION SET

We are only interested in the here-and-now decision to be taken. Thus, a solution to (1) is not a pair (x^o, y^o). To see this, it suffices to remark that once x is selected and ξ is observed, the set of optimal second stage decisions y, is uniquely determined by solving the linear program:

$$\text{(2)} \qquad \begin{array}{ll} \text{Minimize} & qy \\ \text{subject to} & My = \xi - Tx \\ & y \geq 0 \end{array}$$

It is thus obvious that the only decision variable of problem (1) is x.

Nevertheless, the second stage affects our decision on x in two ways. First, we need to limit our set of acceptable first stage decision to those for which there exists a feasible second stage decision, i.e. problem (2) is feasible. Also, for each selection of a vector x, we must take into account the expected costs of the second stage decisions such an x may generate: $E_\xi\{\text{Min } qy \mid x\}$.

A. The Set of Feasible Solutions

A feasible solution to (1) is a vector x such that it satisfies the first stage constraints and such that it is always possible to find a feasible solution to the second stage problem (2), whatever be the value assumed by ξ on Ξ. Dantzig and Madansky [2] call such a solution a permanently feasible solution. The word "permanently" was introduced to reinforce this notion of feasibility of the second stage problem for all values of ξ. We have rejected this terminology because it sometimes leads to confusion in the understanding of problem (1).

The following example shows how the Dantzig - Madansky definition of permanently feasibility differs from what one believes is meant by permanently feasibility. We reserve the terms "permanently feasible" for the following concept: Select a vector x such that the constraints be satisfied with probability one. Consider the following problem:

$$(3) \qquad \text{Minimize } \bar{z}(x) = cx + Q(Tx - \xi) \quad x \in \Omega$$

where ξ is a $\bar{m}$-dimensional random vector on $(\Xi, \mathcal{F}, F)$, T is a matrix $\bar{m} \times n$, $\Omega = \{x | Ax = b,\ x \geq 0\} \subset \mathbb{R}^n$ and Q is a real-valued function. If Q is defined as follows:

$$Q(Tx - \xi) = 0 \qquad \text{if } Tx \geq \xi$$

$$Q(Tx - \xi) = +\infty \qquad \text{otherwise}$$

then, for each given ξ, (3) is a linear programming problem. Such a function $Q(Tx - \xi)$ requires permanent feasibility, i.e. if there exists a solution $(\bar{z}(x) \neq +\infty)$ to problem (3) it must satisfy the condition:

$$(4) \qquad Tx \geq \xi \qquad \forall \xi \in \Xi$$

To see that problem (1) is not as restrictive, e.g. let

$$Q(Tx - \xi) = E_{\xi}\left\{\sum_{i=1}^{\bar{m}} Q_i(T_i x - \xi_i)\right\}$$

where

$$Q_1(T_1x - \xi_1) = 0 \qquad \text{if } T_1x \geq \xi_1$$
$$Q_1(T_1x - \xi_1) = q_1^-(\xi_1 - T_1x) \qquad \text{if } T_1x \leq \xi_1$$

Such a function $Q(Tx - \xi)$ does no longer impose permanently feasibility, i.e. $\bar{z}(x)$ is no longer identically equal to $+\infty$ for all x which does not satisfy condition (4). We can then rewrite (3) as follows:

$$(5) \qquad \begin{array}{lll} \text{Minimize} & z(x) = cx + E_\xi\{0.y^+ + q^-y^-\} & \\ \text{subject to} & Ax & = b, \\ & Tx + \qquad Iy^+ - Iy^- = \xi, & \xi \text{ on } (\Xi, \mathcal{F}, F) \\ & x \geq 0, \qquad y^+ \geq 0,\ y^- \geq 0 & \end{array}$$

Problem (5) is a special case of problem (1), known as the complete problem [5].

From our definition of feasible solution, it is clear that the decision maker is limited in its decision by a double set of constraints. Let

$$K_1 = \{x | Ax = b, \quad x \geq 0\}.$$

We say that K_1 is the set determined by the <u>fixed constraints</u>.

(6) <u>Proposition</u>: K_1 is a convex polyhedron

A set <u>C is convex</u> if $x_1, x_2 \in C$ implies $[x_1, x_2] \subset C$. By <u>convex polyhedron</u> we mean that K_1 can be written as the sum of a convex polytope (convex hull of a finite number of points in $\mathfrak{R}^n$) and a convex polyhedral cone.

Let

$$K_2 = \{x \mid \forall \xi \in \Xi,\ \exists y \geq 0 \text{ such that } My = \xi - Tx\}$$

We say that K_2 is the set representing the constraints imposed on our vector x by the induced constraints. The word "induced" means that these constraints are the restrictions imposed on x by the condition: The second stage problem (2) must be feasible for all $\xi \in \Xi$. This is the real meaning of the equality sign found in the constraints of the standard form:

$$Tx + My = \xi, \qquad \xi \text{ on } (\Xi, \mathcal{G}, F)$$

Let

$$K_{2\xi} = \{x \mid Tx = \xi - My \text{ for some } y \geq 0\}$$

It is easy to see that $K_{2\xi}$ is a convex polyhedron.

(7) Proposition: K_2 is convex

We have $K_2 = \bigcap_{\xi \in \Xi} K_{2\xi}$, then K_2 is either empty, a singleton or for all pairs of points $x_1, x_2 \in K_2$ we have $x_1, x_2 \in K_{2\xi}$ for all $\xi \in \Xi$. Then $\forall \xi \in \Xi$, $[x_1, x_2] \subset K_{2\xi}$, also $[x_1, x_2] \subset \bigcap_{\xi \in \Xi} K_{2\xi} = K_2$.

Obviously,

(8) Proposition: $K = K_1 \cap K_2$ is a convex set

where K is the set of feasible solutions. Remark that we have expressed the set of feasible solution in terms of x along, rather than x and y.

In what follows, we assume that K has full dimension. If this were not the case, one would need to appeal to the relative topology. Most of our proofs do not require this assumption, but it simplifies our treatment and terminology.

The set K_1 is immediately available in terms of linear equations and inequalities involving x only. The set K_2 presents much more difficulty. In general, say when Ξ is a continuum, i.e. $\bigcap_{\xi \in \Xi} K_{2\xi}$ is an infinite intersection of convex polyhedrons, then the characterization of K_2 in terms of x along is a much more complex problem. One main difficulty one encounters in trying to solve a program under uncertainty (no assumptions on the probability space or on the structure of the constraints of (1)) lies in determining whether or not a given x belongs to K.

We now examine some special case where the assumptions made either on the constraints structure of problem (2) or on the probability space $(\Xi, \mathcal{F}, F)$ allow us to obtain fairly easily an explicit expression for the set K_2 (and so for K).

1. <u>Ξ has a finite number of points (Card $|\Xi| < \infty$)</u>

The intersection $\bigcap_{\xi \in \Xi}$ is finite and since $K_{2\xi}$ is a convex polyhedron, so is K_2, and so is K. Let $\xi^1, \xi^2, \ldots, \xi^k$ be the values of ξ for which $f(\xi) \neq 0$. Then,

$$K_2 = \{x \mid Tx + My^{\ell} = \xi^{\ell},\ \ell = 1, \ldots, k\}$$

2. The matrix $M = I$ (identity) and Ξ is compact

Then

$$K_2 = \{x \mid \forall \xi \in \Xi, \exists y \geq 0 \text{ and } y = \xi - Tx\}$$

which implies

$$x \in K_2 \text{ iff } \forall \xi \in \Xi, \xi - Tx \geq 0$$

Since Ξ is bounded, $\exists$ a smallest closed interval, say $\Xi^* \subset \Re^{\bar{m}}$, with lower bound α, such that $\Xi \subset \Xi^*$. The α_i's correspond to the lower bounds for the random variables ξ_i, $i = 1, \ldots, \bar{m}$.

(9) Proposition: $\forall \xi \in \Xi, \xi - Tx \geq 0$ iff $Tx \leq \alpha$

The proof of this proposition is trivial. We have,

(10) $$K_2 = \{x \mid Tx \leq \alpha\}$$

3. $M = (I, -I)$. The problem is complete

One says that problem (1) is complete [5] when the matrix M (after an appropriate rearrangement of rows and columns) can be partioned in two parts, whose first part is the identity matrix and the second part is the negative of an identity matrix, $M = (I, -I)$. This case seems to represent a very important class of applications of programming under uncertainty. It is thus an encouraging fact that the set K can be expressed immediately in terms of linear constraints in x. No assumption at all is necessary on the probability space $(\Xi, \mathcal{F}, F)$.

Let us partition the vector y as follows

$$y = (y^+, y^-)$$

where y^+ corresponds to I and y^- to $-I$, then

$$K_2 = \{x \mid \forall \xi \in \Xi, \exists y^+ \geq 0,\ y^- \geq 0 \text{ such that } y^+ - y^- = \xi - Tx\}$$

(11) <u>Proposition</u>: $K = K_1$

Since $K_2 = \mathfrak{R}^n$ (it is always possible to express any number as the difference of two non-negative numbers), we have $K = K_1 \cap \mathfrak{R}^n = K_1$.

This property, $K = K_1$, gives an intuitive justification for the use of the word "complete". Nevertheless, we should remark that $K = K_1$ does not imply that $M = (I, -I)$.

B. <u>A Feasibility Test</u>

We now fix x and ξ and concentrate our attention on the feasibility of problem (2). From Farkas' lemma we get:

(12) Either the equations

$$My = \xi - Tx$$

have a non-negative solution or the inequalities

$$uM \geq 0 \qquad u(\xi - Tx) < 0$$

have a solution.

(13) <u>Propositon</u>: $x \in K_2$ iff $\forall \xi \in \Xi$ we have $U(x,\xi) \geq 0$, where

$U(x,\xi) = \{\text{Min } u(\xi - Tx) \mid uM \geq 0\}$

If for a given $\bar{x}$ and $\forall \xi \in \Xi$ we have $U(\bar{x},\xi) \geq 0$, it implies that the system of inequalities $uM \geq 0$ and $u(\xi - Tx) < 0$ have no solution. By (12), the system $My = \xi - T\bar{x}$ has then a non-negative solution, for all $\xi \in \Xi$. This means that $\bar{x} \in K_2$.

Proposition (13) yields a test which allows us to determine if a given $x \in K_1$, is or is not a feasible solution to (1). Nonetheless, such a procedure would be completely inefficient if we had to perform this test for all ξ in Ξ. If Ξ does not have finite cardinality, this test for any given x would involve solving an infinite number of linear programs of the form:

$$\begin{aligned} &\text{Minimize} && u(\xi - Tx) \\ &\text{subject to} && uM \geq 0 \end{aligned}$$

If problem (1) is stated in a slightly different form (it is very often possible to reduce problem (1) to (14)), viz.:

$$\text{(14)} \qquad \begin{aligned} &\text{Minimize} && z(x) = cx + E_{\xi}\{qy\} \\ &\text{subject to} && Ax = b \\ & && Tx + My \geq \xi, \qquad \xi \text{ on } (\Xi, \mathcal{F}, F) \\ & && x \geq 0, \quad y \geq 0 \end{aligned}$$

it is possible to obtain a more efficient test. We then apply the following form of Farkas' lemma: Exactly one of the two alternations hold: Either the inequality

$$My \geq \xi - Tx$$

has a non-negative solution, or the inequalities

$$uM \geq 0 \qquad u(\xi - Tx) < 0$$

have a non-negative solution.

Then

(15) <u>Proposition</u>: $x \in K$ iff $x \in K_1$ and $\forall \xi \in \Xi$, $U(x,\xi) \geq 0$, where

$U(x,\xi) = \{\text{Min } u(\xi - Tx) \mid uM \geq 0, u \geq 0\}$

If Ξ has a lower bound--from a practical point of view this is a very mild condition--then, let α be such that $\alpha \in \Xi$ and $\alpha_i \leq \xi_i$ for all $\xi_i \in \Xi_i$, $i = 1,\ldots,\bar{m}$. Since u is restricted to be non-negative, we have

$$U(x,\alpha) \leq U(x,\xi) \qquad \forall \xi \in \Xi$$

Moreover, $\alpha \in \Xi$ and $U(x,\alpha) < 0$ imply that there exists at least one point of Ξ for which the condition $U(x,\xi) \geq 0$ does not hold. By Proposition (10) this x is not a feasible solution. We have proved:

(16) <u>Proposition</u>: $x \in K$ iff $x \in K_1$ and $U(x,\alpha) \geq 0$

For this case, it is thus sufficient to solve one linear program to test the feasibility of a given x which belongs to K_1. Proposition (11) is not true if $\alpha_i \leq \xi_i$ for all $\xi_i \in \Xi_i$, $i = 1,\ldots,\bar{m}$ but $\alpha \notin \Xi$. For instance, consider the following example:

Let

$$T = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad M = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$$

$$\Xi = \{\xi \mid -1 \leq \xi_1 \leq 0, \quad \xi_2 \leq 2, \quad \xi_1 + \xi_2 \leq 0\}$$

and let $\bar{x} = (\bar{x}_1, \bar{x}_2) = (0,0)$ belong to K_1. By definition of Ξ, $\alpha = (\alpha_1, \alpha_2) = (-1,0)$. It is easy to see that $\Xi \subset \{\zeta \mid \zeta = My, \ y \geq 0\}$ and that the affine transformation obtained by translating Ξ by Tx, map Ξ into itself $(T\bar{x} = 0)$, i.e. $\bar{x} = 0 \in K$. But $U(\bar{x},\alpha)$ is not bounded below.

Suppose now that we have at hand $\hat{x}$ such that $\hat{x} \in K_1$ and $U(\hat{x},\alpha) < 0$, where $U(x,\alpha)$ is as defined in (15). Let $\hat{u}$ be an optimal solution to:

$$\begin{aligned} &\text{Minimize} && u(\alpha - T\hat{x}) \\ &\text{subject to} && uM \geq 0 \\ & && u \geq 0 \quad . \end{aligned}$$

Since $U(\hat{x},\alpha) < 0$, we have $\hat{u}\alpha < \hat{u}T\hat{x}$ and by (16) $\hat{x} \notin K$. Thus, every $x \in K$ must satisfy the inequality:

(17) $$(\hat{u}T)x \leq \hat{u}\alpha \ .$$

We can add this condition (12) to the fixed constraints, $Ax = b$, $x \geq 0$. It has the effect of <u>cutting off</u> part of the set K_1.

III. THE EQUIVALENT CONVEX PROGRAMMING PROBLEM

We now show that a linear program under uncertainty can be expressed in terms of first stage decision variable x, as a convex program that we shall call the equivalent convex programming problem. We derive the properties of the objective function of the equivalent convex program and construct the equivalent convex program when the constraints and the probability space satisfy the assumptions made in Section II.

A. The Equivalent Convex Program

(18) Definition: A programming problem: Minimize $f(x)$, $x \in K$, is an equivalent programming problem to (1), if $f(x)$ is given explicitly for each x (not just as a function of x, y, and ξ as in (1')), if K is the set of feasible solutions to (1), and if an optimal solution to the equivalent programming problem is an optimal solution to (1).

In Section II, we have already characterized the set of feasible solutions to (1). To exhibit an equivalent convex program to (1), it suffices to show that (1') is convex in x. Let us consider the second stage problem (2) for a fixed ξ in Ξ, as a function of x. Then, by Proposition (3) of Appendix I,

$$P(x,\xi) = \{\text{Min } qy \mid My = \xi - Tx,\ y \geq 0\} \tag{19}$$

is a convex of x on $\{x \mid Tx = \xi - My,\ y \geq 0\}$ and in particular on K_2.

By the duality theorem for linear programs, we have

(20) $$P(x,\xi) = Q(x,\xi)$$

where

$$Q(x,\xi) = \{\text{Max } \pi(\xi - Tx) \,|\, \pi M \leq q\} \qquad \text{for fixed } \xi \text{ in } \Xi.$$

Let

(21) $$Q(x) = E_\xi\{\text{Min } qy \,|\, My = \xi - Tx,\ y \geq 0\} = E_\xi\{Q(x,\xi)\} = E_\xi\{P(x,\xi)\}$$

be the expected value of the second stage problem (2) for a given x in K_2.

(22) <u>Proposition</u>: $Q(x)$ is convex on K_2 [2].

Since by Proposition (3) of Appendix I, $Q(x,\xi)$ is convex in x on K_2, it suffices to remark that applying the operator E_ξ to $Q(x,\xi)$ is equivalent to performing a positive weighted linear combination of convex functions, i.e. $Q(x)$ is convex on K_2.

Thus the equivalent convex program to (1) is,

(23)

Minimize	$z(x) = cx + Q(x)$
subject to	$x \epsilon K$

(24) <u>Proposition</u>: $Q(x)$ is continuous on K_2.

Since $Q(x)$ is convex on K_2, the result is immediate if K_2 is open. To see that K_2 could be open, consider the following example:

Let $M=1$, $T=1$ and $\Xi=(0,1)$, then $K_2=(-\infty, 0)$. In general, by Corollary (12) of Appendix I, $Q(x,\xi)$ is uniformly continuous in x and ξ, thus $Q(x) = \int_{\xi\in\Xi} Q(x,\xi)\,dF(\xi)$ is continuous in x on K_2.

Consider the dual to the second stage problem (2),

(25) $$\text{Maximize} \quad \pi(\xi - Tx)$$
$$\text{subject to} \quad \pi M \leq q,$$

and let $\pi(x,\xi)$ be the optimal solution to (25) for fixed x and ξ. In what follows, we assume that $\pi(x,\xi)$, and $Q(x,\xi)$ are defined for all x in K, and all ξ in Ξ. Define

(26) $$\pi(x) = E_\xi\{\pi(x,\xi)\} = \int_{\xi\in\Xi} \pi(x,\xi)\,dF(\xi)$$

as the expected optimal solution to problem (25) for a given x. Also, let

$$\psi(x) = E_\xi\{\pi(x,\xi)\xi\} = \int_{\xi\in\Xi} \pi(x,\xi)\,\xi\,dF(\xi).$$

Note that $\pi(x)$ is a $\bar{m}$-dimensional vector and that $\psi(x)$ is a scalar.

(27) <u>Proposition</u>: $[c - \pi(\bar{x})T]x = -\psi(\bar{x})$ is a supporting hyperplane of $z(x)$ at $x=\bar{x}$, where $\bar{x}\in K$ [2].

Since $[c - \pi(\bar{x})T]\bar{x} + \psi(\bar{x}) = z(\bar{x})$, it suffices to show that, $\forall x\in K$, $z(x) \geq [c - \pi(\bar{x})T]x + \psi(\bar{x})$. But this is true, since for all $x\in K$, and for all ξ in Ξ, $\pi(x,\xi)(\xi - Tx) \geq \pi(\bar{x},\xi)(\xi - Tx)$. Integrating both sides with respect to $dF(\xi)$ and adding cx on both sides, we get:

$$z(x) = [c - \pi(x)T]x + \psi(x) \geq [c - \pi(\bar{x})T]x + \psi(\bar{x}).$$

(28) <u>Corollary</u>: $[c - \pi(\bar{x})T]$ is a gradient of $z(x)$ at $\bar{x}$.

We use the term "gradient", as defined in (7) of Appendix II.

(29) <u>Proposition</u>: If $F(\xi)$ is continuous, then $z(x)$ is differentiable on K.

By Proposition (15) of Appendix I, $\pi(x,\xi)$ is piece-wise constant. Moreover, since the set of points where $\pi(x,\xi)$ is multi-valued has measured zero, $\pi(x)$ and $\psi(x)$ are unique for all $x \in K$. This implies that $z(x)$ has z unique supporting hyperplane for all x in K. By Proposition (11) of Appendix II, $z(x)$ is differentiable on K.

The condition $F(\xi)$ continuous, is sufficient but not necessary, e.g., let

$$T = M = I, \quad \Xi = \{\xi^1 = \binom{1}{2}, \quad \xi^2 = \binom{2}{1}\}, \quad f(\xi^1) = f(\xi^2) = 1/2$$

$$c = [2,2], \quad q = [1,1], \quad x = [x_1, x_2]$$

then $z(x) = x_1 + x_2$.

(30) <u>Proposition</u>: Let $x^o \in K$, then x^o is optimal iff $\exists\, \pi(x^o)$ such that $\forall x \in K$, $[c - \pi(x^o)T]x^o \leq [c - \pi(x^o)T]x$.

The proof is a direct application of (28) and Proposition (16) of Appendix II.

(31) <u>Corollary</u>: If $z(x)$ is differentiable, then x^o is optimal iff for all x in K, $[c - \pi(x^o)T]x^o \leq [c - \pi(x^o)T]x$.

One could regard (30) and (31) as statements related to the solvability of problem (1). If we disregard the inconsistent case (K is empty), we can write: (1) is solvable iff $\exists$ a pair $(x^o, \pi(x^o))$ such that $[c - \pi(x^o)T]x^o \leq [c - \pi(x^o)T]x$, for all x in K. Note that (1) can have an infinite or a finite infimum. Moreover, since $z(x)$ is continuous, $z(x)$ may fail to achieve a minimum on K only if K is not bounded.

B. <u>Special Cases</u>

When the constraints of problem (1) and the probability space satisfy the assumptions considered in Section I, we show that the equivalent convex programs are programming problems for which satisfactory algorithms exist.

1. <u>Ξ is finite</u>

Let $\xi^1, \xi^2, \ldots, \xi^k$ be the values assumed by the random vector ξ with probabilities $f^1, f^2, \ldots, f^k$, respectively. We have seen in Section I that the induced constraints can be expressed explicitly in terms of linear equations and linear inequalities. The equivalent convex program is a linear programming problem which can be expressed as follows:

$$(32) \qquad \begin{array}{llllll} \text{Minimize} & z(x) = cx + f^1qy^1 + f^2qy^2 + \ldots + f^kqy^k & & & & \\ \text{subject to} & Ax & & & & = b \\ & Tx + & My^1 & & & = \xi^1 \\ & Tx + & & My^2 & & = \xi^2 \\ & \vdots & & & \ddots & \vdots \\ & Tx + & & & My^k & = \xi^k \\ & x \geq 0,\ y^1 \geq 0,\ y^2 \geq 0, \ldots, y^k \geq 0. & & & & \end{array}$$

Dantzig and Madansky [2] have shown that there exist a dual of this problem which is in the standard form for the application of the decomposition algorithm of Dantzig and Wolfe [4]. To find this dual problem, we use a more direct approach than the one found in [2].

Let $(\sigma, \bar{\pi}^1, \bar{\pi}^2, \ldots, \bar{\pi}^k)$ be the variable appearing in the usual dual formulation of (32). Define

$$\pi^\ell = \frac{1}{f^\ell}\bar{\pi}^\ell \qquad \ell = 1, \ldots, k,$$

then the dual reads:

$$\begin{array}{llllll} \text{Maximize} & \sigma b + f^1\pi^1\xi^1 & + f^2\pi^2\xi^2 & + \ldots + & f^k\pi^k\xi^k & \\ & \sigma A + f^1\pi^1 T & + f^2\pi^2 T & + \ldots + & f^k\pi^k T & \leq c \\ & \pi^1 M & & & & \leq q \\ & & \pi^2 M & & & \leq q \\ & & & \ddots & & \vdots \\ & & & & \pi^k M & \leq q \end{array}$$

This problem has an "angular" structure. The first n inequalities can be used to generate the master program. The last $k \times \bar{n}$ inequalities constitute the sub-problem. Depending on T and M, it may be advantageous to use variants of the decomposition algorithm, e.g., see Abadie [1].

Another simple transformation gives the problem (32) the structure of a multi-stage system (so called "staircase" system) where the linear constraints for all stages but one are identical. This last feature may simplify considerably the computation. To obtain this form, subtract from each row of $Tx + My^{\ell+1} = \xi^{\ell+1}$ the corresponding row of $Tx + My^{\ell} = \xi^{\ell}$ for $\ell = 1, \ldots, k-1$. Problem (32) becomes:

$$
\begin{array}{lllllll}
\text{Minimize} & z(x) = cx + f^1 qy^1 + f^2 qy^2 + \ldots + f^{k-1} qy^{k-1} + f^k qy^k & & & & & \\
\text{subject to} & Ax & & & & & = b \\
& Tx + & My^1 & & & & = \xi^1 \\
& - & My^1 + & My^2 & & & = \xi^2 - \xi^1 \\
& & & \cdot & \cdot & \cdot -My^{k-1} + My^k & = \xi^k - \xi^{k-1} \\
& x \geq 0, & y^1 \geq 0, & y^2 \geq 0, & \cdots & y^{k-1} \geq 0, \ y^k \geq 0 . &
\end{array}
$$

2. M is square, non-singular and Ξ is bounded

We show that under these assumptions there exists a linear programming problem whose set of optimal solutions is the set of optimal solutions of the linear program under uncertainty.

a) M is the identity (M = I)

The problem under consideration is:

$$\begin{aligned} \text{Minimize} \quad & z(x) = cx + E_\xi\{qy\} \\ \text{subject to} \quad & Ax = b \\ & Tx + Iy = \xi, \quad \xi \text{ on } (\Xi, \mathcal{F}, F) \\ & x \geq 0, \quad y \geq 0 \end{aligned}$$

For fixed x and ξ, the second stage problem (2) is:

(33)
$$\begin{aligned} \text{Minimize} \quad & qy \\ \text{subject to} \quad & Iy = \xi - Tx \\ & y \geq 0 \end{aligned}$$

If (33) is feasible, then $\text{Min } qy = q(\xi - Tx)$. Moreover, if $x \in K$, then (33) is feasible for all ξ in Ξ, i.e. $\xi - Tx \geq 0$ for all ξ in Ξ. We have

$$E_\xi\{\text{Min } qy \mid x \in K\} = E_\xi\{q(\xi - Tx)\} = q\bar{\xi} - qTx$$

By (9) and (10) there exists a vector α such that

$$K = \{x \mid Ax = b, \quad Tx \leq \alpha, \quad x \geq 0\}$$

Thus, the linear program

$$\begin{aligned} \text{Minimize} \quad & z(x) = (c - qT)x \\ \text{subject to} \quad & Ax = b \\ & Tx \leq \alpha \\ & x \geq 0 \end{aligned}$$

yields the set of optimal solutions to our problem. If Ξ is compact, then each neighborhood of α_i has positive measure. If the random variables ξ_i, $i=1,\ldots,\bar{m}$, are independent, then Ξ is an interval (in $\mathfrak{R}^{\bar{m}}$) and $\Xi^* = \Xi$ (9).

b) <u>M is square and non-singular</u>

The problem reads:

$$\begin{array}{lll} \text{Minimize} & z(x) = cx + E_\xi\{qy\} & \\ \text{subject to} & Ax \qquad\qquad = b, & \\ & Tx + \quad My = \xi, & \xi \text{ on } (\Xi, \mathcal{G}, F) \\ & x \geq 0, \quad y \geq 0 \,. & \end{array}$$

If one multiplies to the left, both sides of $Tx + My = \xi$ by $M^{-1} = [\mu_{ij}]$ one obtains:

$$\tilde{T}x + Iy = \tilde{\xi}, \qquad \tilde{\xi} \text{ on } (\tilde{\Xi}, \tilde{\mathcal{G}}, \tilde{F})$$

where

$$\tilde{T} = M^{-1}T$$
$$M^{-1}: \Xi \to \tilde{\Xi}$$

Since Ξ is bounded by assumption and M^{-1} is a non-singular linear mapping, $\tilde{\Xi}$ is also bounded. Hence, our new problem is similar to the previous case $(M=I)$. Let Ξ^* be the smallest interval containing $\tilde{\Xi}$ and let α^* be the lower bound of Ξ^*. The equivalent convex program then reads:

$$\text{Minimize} \quad z(x) = (c - qM^{-1}T)x$$
$$\text{subject to} \quad Ax = b$$
$$M^{-1}Tx \leq \alpha^*$$
$$x \geq 0$$

The components of the vector α^* can be computed as follows: Let α_i and β_i be respectively the greatest lower bound and the least upper bound for ξ_i, then

$$\alpha_i^* = \text{Min} \sum_{j=1}^{\bar{m}} \mu_{ij}\xi_j \quad \text{where} \quad \alpha_j \leq \xi_j \leq \beta_j$$

$$= \sum_{j=1}^{\bar{m}} \mu_{ij}\xi_j^* \quad \text{where} \quad \xi_j^* = \alpha_j \text{ if } \mu_{ij} \geq 0$$
$$\xi_j^* = \beta_j \text{ if } \mu_{ij} < 0$$

From this computation procedure for α^*, it is easy to see that the condition, Ξ bounded, is too strong; all we need is that α^* exists.

Some generalizations are possible, for example: let M be a Leontief matrix with substitutions and let $\xi - Tx \geq 0$ for all ξ in Ξ and all x in K, one then shows that such a problem can be reduced to the case where M is square and non-singular [3]. In this case, the condition $\xi - Tx \geq 0$, for all ξ in Ξ, and all x in K is not restrictive, since $\xi_i - T_i x < 0$, for some i, is meaningless if the second stage problem (2) is a Leontief system with substitution.

3. The problem is complete. $M = (I, -I)$

By Proposition (11), the equivalent convex program has the form:

$$\text{(34)} \qquad \begin{array}{ll} \text{Minimize} & z(x) = cx + Q(x) \\ \text{subject to} & Ax = b \\ & x \geq 0 \end{array}$$

This problem was studied in detail in [5]. For completeness, we list the particular forms of this convex program for some specific distribution functions, $F(\xi)$.

Assumptions on $(\Xi, \mathcal{G}, F)$	Equivalent Convex Program
Ξ is finite (ξ discrete)	Linear program with upper bounds
$F(\xi)$ uniform	Quadratic Program
$F(\xi)$ continuous	
If one approximates ξ by a sum of uniformly distributed random variables, then	Quadratic Program
$F(\xi)$ exponential	
If one approximates the objective function, then	Quadratic Program
In general	Separable Convex Program

Moreover, many generalizations of the complete problem lead to an identical class of equivalent convex programs. Let us, for instance, consider the following problem:

$$\begin{array}{ll} \text{Minimize} & z(x) = cx + E_{\xi}\{q^{+}y^{+} + q^{-}y^{-}\} \\ \text{subject to} & Ax = b \\ & Tx + Iy^{+} - Iy^{-} = \xi, \ \xi \text{ on } (\Xi, \mathcal{G}, F) \\ & x \geq 0, \quad y^{+} \in H, \quad y^{-} \geq 0 \end{array}$$

where

$$H = \{y^+ | y^+ = Lz, \ z \geq 0\}$$

If L is a Leontief matrix with substitution such that H contains some $y^+ > 0$, and if $q^+ + q^- \geq 0$, then one can show [3] that such a problem has also an equivalent convex program of the form (34).

REFERENCES

[1] J. Abadie, "On Decomposition Principle", Operations Research Center, University of California, Berkeley, ORC 63-20, August, 1963.

[2] G. B. Dantzig and A. Madansky, "On the Solution of Two-Stage Linear Programs under Uncertainty", Proceedings of the Fourth Symposium on Mathematical Statistics and Probability, Vol. I, University of California, Berkeley, 1961.

[3] G. B. Dantzig and R. Wets, "Leontief Matrices with Substitutions", Operations Research Center, University of California, Berkeley, ORC 63-19, 1963.

[4] G.B. Dantzig and P. Wolfe, "Decomposition Principle for Linear Programs", Operations Research, Vol. 8, No. 1.

[5] R. Wets, "Programming under Uncertainty: The Complete Problem", Boeing Scientific Research Laboratories, Seattle, D1-82-0379, October, 1964.

Appendix I

A linear program can be considered as a function of its parameters

$$f(c,A,b) = \{\text{Min } cx \mid Ax = b, \quad x \geq 0\}.$$

We study the properties of this function where b is variable. Let

$$f(t) = \{\text{Min } cx \mid Ax = t, \quad x \geq 0\}$$

$$\mathcal{C} = \{t \mid t = Ax, \quad x \geq 0\}$$

(1) Lemma: $\mathcal{C}$ is a convex polyhedral cone containing the origin, 0. For the sake of simplicity, we shall assume that the matrix A has full row rank, so in particular $m \leq n$. The case $f(t) = -\infty$ for all t in $\mathcal{C}$ is without interest, and since

(2) Lemma: $f(t) = -\infty$ for some $t \in \mathcal{C}$ iff $\forall t \in \mathcal{C}$, $f(t) \equiv -\infty$, we shall assume in what follows that $f(t) > -\infty$ for all $t \in \mathcal{C}$. Note that $f(t)$ is defined only for t in $\mathcal{C}$.

(3) Proposition: $f(t)$ is convex on $\mathcal{C}$.

Consider any $t_0, t_1 \in \mathcal{C}$ and $\lambda \in [0,1]$. Let $t_\lambda = \lambda t_0 + (1-\lambda)t_1$, by (1) we have, $t_\lambda \in \mathcal{C}$. Let x_i be such that

$$f(t_i) = cx_i = \{\text{Min } cx \mid Ax = t_i, \quad x \geq 0\} \quad \text{for} \quad i = 0,\lambda,1$$

then $\bar{x} = \lambda x_0 + (1-\lambda)x_1$ is a feasible but not necessarily optimal solution to: Min cx such that $Ax = t_\lambda$, $x \geq 0$. Consequently, we have that $f(t)$ satisfies the Jensen's inequality,

$$\lambda f(t_0) + (1-\lambda)f(t_1) = \lambda c x_0 + (1-\lambda)c x_1 = c\bar{x} \geq c x_\lambda = f(t_\lambda)$$

for all t_0, t_1 in $\mathfrak{C}$ and $0 \leq \lambda \leq 1$. Loosely speaking, we can rephrase (3) as follows: A linear program is a convex function of its right-hand side.

(4) Corollary: Let $f^*(t) = \{\text{Min } tx \mid Ax = b,\ x \geq 0\}$ and let $\mathfrak{C}^* = \{t \mid f^*(t) > -\infty\}$. Then $f^*(-t)$ is concave on $\mathfrak{C}^*$.

(5) Proposition: If A is square and non-singular, then $\mathfrak{C}$ is a simplicial cone and $f(t)$ is linear on $\mathfrak{C}$.

It suffices to remark that $f(t) = cA^{-1}t$ on $\mathfrak{C} = \{t \mid A^{-1}t \geq 0\}$.

(6) Proposition: Let B be a submatrix of A such that B is an optimal basis for some t. Then B is an optimal basis for all t in $\mathfrak{C}_B = \{t \mid B^{-1}t \geq 0\}$, $_B$ is a simplicial cone, and $\mathfrak{C}_B \subset \mathfrak{C}$.

(7) Corollary: If B is an optimal basis for some t, then $\mathfrak{C}_B$ is the unique subset of $\mathfrak{C}$ for which B constitutes an optimal basis.

By Propositions (5) and (6):

(8) Corollary: $f(t)$ is linear on $\mathfrak{C}_B$.

(9) Proposition: There exists a decomposition of $\mathfrak{C}$ into simplicial cones $\mathfrak{C}_1, \ldots, \mathfrak{C}_k$ such that

(i) $\mathcal{C}_i = \{t \mid B_i t \geq 0\}$ $i = 1, \ldots, k$, where B_i is a square, non-singular matrix of A of rank m,

(ii) B_i is an optimal basis for some t,

(iii) $\bigcup_{i=1}^{k} \mathcal{C}_i = \mathcal{C}$,

(iv) $\text{int}\, \mathcal{C}_i \cap \text{int}\, \mathcal{C}_j = \emptyset$ for $i \neq j$.

This proposition can be proved using (5), (6) and (7). It is easy to see that this decomposition may not be unique. By Propositions (8) and (9) we get:

(10) <u>Proposition:</u> $f(t)$ is piece-wise linear on $\mathcal{C}$.

(11) <u>Proposition:</u> $f(t)$ is continuous on $\mathcal{C}$.

Since $f(t)$ is convex it is continuous on the $\text{int}\, \mathcal{C}$. Moreover, (8), (9) and (10) imply that $f(t)$ is linear on, and in the neighborhood of, the boundary.

(12) <u>Corollary:</u> $f(t)$ is uniformly continuous on $\mathcal{C}$.

This is immediate by (10) and (11).

Consider the following problem:

(13) Maximize πt subject to $\pi A \leq c$

and let $\pi(t)$ be an optimal solution to (13) for a given t in $\mathcal{C}$.

(14) <u>Proposition:</u> If A is square and non-singular, then $\pi(t)$ is constant on $\mathcal{C}$.

It suffices to remark that $\pi(t) = cA^{-1}$ on $\mathcal{C} = \{t \mid A^{-1} t \geq 0\}$.

(15) <u>Proposition:</u> $\pi(t)$ is a piece-wise constant function on $\mathbf{C}$.

This proposition can be proved using (9) and (14). Let us remark that $\pi(t)$ may be multi-valued on the boundaries of the simplicial cones determining the decomposition of $\mathbf{C}$, but it is single valued on their interior.

(16) <u>Proposition:</u> $\pi(\bar{t})\cdot t$ is a supporting hyperplane to $f(t)$ at $t = \bar{t}$, $\bar{t}\in\mathbf{C}$.

Since at $t=\bar{t}$, the hyperplane $\pi(\bar{t})\cdot t$ intersects $f(t)$, it suffices to show that,

$$\pi(\bar{t})\cdot t \leq f(t) \qquad\qquad \forall t\in\mathbf{C}.$$

But this is true, since by the definition of $\pi(t)$,

$$\pi(\bar{t})\cdot t \leq \pi(t)\cdot t = f(t).$$

This last proposition, (10) and (11) imply that

(17) <u>Proposition:</u> The graph of $f(t)$, $\{(z,t)\,|\,z \geq f(t),\ t\in\mathbf{C}\}$, is a convex polyhedral cone with vertex 0.

Appendix II

The purpose of this appendix is to review some pertinent properties of convex functions, define a "gradient" at all points of the domain of a convex function, and establish a duality theorem for the following problem:

(1) $$\text{Minimize } f(x), \text{ subject to: } Ax = b,\ x \geq 0$$

In what follows, we assume that all sets considered are of full dimension; if this was not, one would need to appeal to the relative topology. Since the proofs of the following propositions are standard, we have either limited ourselves to an outline, or left the proof altogether to the reader.

A. Convex Sets

A set C is convex iff $x_1, x_2 \in C$ implies $[x_1, x_2] \subset C$. Let $\mathfrak{R}^n = X$ and let X^* denote the conjugate space of X (the set of all continuous linear functionals on X). An element of X^* is denoted by x^* and the origin by 0^*. A hyperplane $x^*(x) = \alpha$ is a bounding hyperplane of C if $\forall x \in C$ we have $x^*(x) \geq \alpha$ and $x^* \neq 0^*$. A bounding hyperplane $x_1^*(x) = \alpha$ is a supporting hyperplane of C at x_1 if $x_1^*(x_1) = \alpha$. Obviously, x_1 belongs to the boundary of C and for all x in the interior of C, $x^*(x) > \alpha$.

(2) Lemma: C is convex iff $\forall x \in$ boundary of C, $\exists$ a supporting hyperplane of C at x.

B. Convex Functions

Let $K \subset \tilde{K} \subset \mathfrak{R}^n$, where $\tilde{K}$ is open and K is closed and convex, and let $f(x)$ be a real-valued function, with domain $\tilde{K}$. A function $f(x)$ is convex on $\tilde{K}$ if $x_1, x_2 \epsilon \tilde{K}$ implies $\lambda f(x_1) + (1-\lambda) f(x_2) \geq f(\lambda x_1 + (1-\lambda) x_2)$, for all λ in $[0,1]$.

(3) Lemma: If $f(x)$ is convex on $\tilde{K}$, then $f(x)$ is continuous on $\tilde{K} (\supset K)$.

(4) Lemma: $f(x)$ is convex on K iff $\{(z,x) | f(x) \leq z,\ x \epsilon K\}$ is a closed, convex subset of $\mathfrak{R}^{n+1}$.

C. "Gradients"

(5) Definition: $s_{\bar{x}}(x) = ax + s_{\bar{x}}(0)$ is a linear supporting function of $f(x)$ at $\bar{x}$ in K if $s_{\bar{x}}(\bar{x}) = f(\bar{x})$ and $s_{\bar{x}}(x) \leq f(x)$ for all x in $\tilde{K}$.

Since all supporting functions to be considered here are linear supporting functions, we shall omit the term linear. Let $S_{\bar{x}} = \{s_{\bar{x}}(x)\}$ be the class of all supporting functions of $f(x)$ at $\bar{x}$ in K.

(6) Lemma: If $f(x)$ is convex, then $S_{\bar{x}} \neq \emptyset$, for all $\bar{x}$ in K.

(4) and (2) imply that there exists a supporting hyperplane to the set $\{(z,x) | f(x) \leq z,\ x \epsilon K\}$ at the point $(f(\bar{x}), \bar{x})$, say $x^*(x) = \alpha$. By continuity of $f(x)$ on $\tilde{K} \subset K$ and since $f(\tilde{K}) \subset \mathfrak{R}$, there exists

a supporting hyperplane at $(f(\bar{x}),\bar{x})$ such that the first component of $x^* = (x_1^*, x_{(n)}^*)$ is different from zero. Thus,

$s_{\bar{x}}(x) = \frac{1}{x_1^*}(-x_{(n)}^* \cdot x + \alpha)$ is a supporting function of $f(x)$ at $\bar{x}$.

(7) Definition: A vector $f_{\bar{x}}$ is a gradient of $f(x)$ at $\bar{x}$ in K, if for some scalar $s_{\bar{x}}(0)$, $s_{\bar{x}}(x) = f_{\bar{x}}x + s_{\bar{x}}(0)$ is a supporting function of $f(x)$ at $\bar{x}$.

By the definition of supporting function, it is easy to see that $s_{\bar{x}}(0) = f(\bar{x}) - f_{\bar{x}} \cdot \bar{x}$. Let $F_{\bar{x}} = \{f_{\bar{x}}\}$ be the class of all gradients of $f(x)$ at $\bar{x}$. Note that there is a one-to-one correspondence between $F_{\bar{x}}$ and $S_{\bar{x}}$. Also, $F_{\bar{x}}$ and $S_{\bar{x}}$ are closed.

(8) Proposition: $f(x)$ is convex iff $\forall x_1, x_2 \epsilon K$ and $\forall f_{x_1}$ in F_{x_1} we have $f(x_2) - f(x_1) \geq f_{x_1}(x_2-x_1)$.

A judicious use of the definition of gradient of (2) and (4) proves this proposition.

(9) Corollary: If $f(x)$ is convex and $f_{x_1}(x-x_1) \geq 0$, for all x in K, where f_{x_1} is a gradient of $f(x)$ at x_1, then $f(x)$ attains its minimum at $x = x_1$.

(10) Corollary: If $f(x^o) \leq f(x)$, for all x in K, then $f_x \cdot x^o \leq f_x \cdot x$.

(11) Proposition: A convex function $f(x)$ is differentiable at $\bar{x}$ in K iff $f_{\bar{x}}$ is unique, i.e. the cardinality of $S_{\bar{x}}$ = card $F_{\bar{x}} = 1$. Then $f_{\bar{x}} = \left(\frac{\partial}{\partial x_1} f, \frac{\partial}{\partial x_2} f, \ldots, \frac{\partial}{\partial x_n} f\right) \Big|_{x = \bar{x}}$.

(12) Proposition: Let $f(x)$ be convex and $x^o \in K$, then $f(x^o) = \{\text{Min } f(x) \mid x \in \tilde{K}\}$ iff $s_{x^o}(x) = f(x^o)$ is a supporting function of $f(x)$ at x^o.

Let us now consider x_1, x_2 in K, let $x_t = tx_1 + (1-t)x_2$ and define $\varphi(t) = f(tx_1 + (1-t)x_2)$ where t belongs to $T = \mathfrak{R} \cap \{t \mid x_t \in \tilde{K}\}$. Since $\tilde{K}$ is open, so is T. We denote a gradient of $\varphi(t)$ at $\bar{t}$, by $\varphi\bar{t}$, and the class of all gradients of $\varphi(t)$ at $\bar{t}$, by $\Phi_{\bar{t}} = \{\varphi_{\bar{t}}\}$. Since $f(x)$ is continuous on $\tilde{K}$, so is $\varphi(t)$ on T. If $f(x)$ is convex, so is $\varphi(t)$.

(13) Proposition: If $f(x)$ is convex, then $\bar{\varphi} \in \Phi_{\bar{t}}$ if $\bar{\varphi} = f_{x_{\bar{t}}}(x_1 - x_2)$ for some $f_{x_{\bar{t}}} \in F_{x_{\bar{t}}}$.

(Hint: Use the monotonicity of the gradient of a convex function, see [2].)

Let $\mathcal{L} = \{x \mid x = tx_1 + (1-t)x_2,\ t \in T,\ x_1, x_2 \in K\}$ and $x^o \in K \cap \mathcal{L}$. In what follows we also assume that $f(x)$ is convex on $\tilde{K}$. The two following propositions are obtained by applying (13) and observing the properties of $\varphi(t)$ on T.

(14) Proposition: If $x^o \in K$ and $f(x^o) \leq f(x)$ for all x in K, then there exists a gradient f_{x_0} such that $f_{x_0} \cdot x_0 = f_{x_0} \cdot x$ for all x in $\mathcal{L}$.

(15) Proposition: Let $x_1, x_2 \in K$, $f_{x_1} \in F_{x_1}$, and $f_{x_1} x_1 > f_{x_1} x_2$, then $\exists\, x^0 \in (x_1, x_2)$ such that $f(x^0) < f(x_1)$.

(16) Proposition: $f(x_0) \leq f(x)$ for all x in K iff $\exists\, f_{x_0} \in F_{x_0}$ such that $f_{x_0} x_0 \leq f_{x_0} \cdot x$ for all x in K.

If $f_{x_0} x_0 \leq f_{x_0} \cdot x$ for all x in K, then by (9) x_0 is optimal. Also, if x is not a minimum, then one can prove that there exists $f_x \in F_x$, such that $f_x \bar{x} < f_x x$ for some $\bar{x} \in K$. Then, (15) yields the proof in the other direction.

(17) Corollary: If $f(x)$ is convex and differentiable, then x_0 is a minimum of $f(x)$ on K iff $f_{x_0} \cdot x_0 \leq f_{x_0} x$ for all x in K, where $f_{x_0} = \left(\frac{\partial}{\partial x_1} f, \ldots, \frac{\partial}{\partial x_n} f \right) \Big|_{x = x_0}$.

D. Special Case: K is a compact, convex polyhedron

Let us now consider problem (1), viz:

(1) Minimize $f(x)$ subject to $x \in K = \{x \mid Ax = b,\ x \geq 0\}$,

where $f(x)$ is a continuous, convex, real-valued function on $\mathfrak{R}^n$. Under these conditions, we show that the dual to (1) reads:

(18)
$$\begin{aligned} &\text{Maximize} \quad f(x) - f_x x + \pi b \\ &\text{subject to} \quad -f_x + \pi A \leq 0 \\ &\qquad\qquad x \geq 0 \end{aligned}$$

where $f_x \epsilon F_x$. Define $K_D = \{(\pi, x) \mid \pi A \leq f_x$ for some $x \geq 0\}$. Obviously, if $x \epsilon K$, then $x \epsilon K_D$.

(19) Weak Duality. For all x in K and (π, x) in K_D, we have $f_x x \geq \pi b$.

To see this multiply the equality constraints of (1) by π and the inequalities of problem (18) by $x \geq 0$.

(20) Strong Duality. There exists $x^o \epsilon K$ and $(\pi^o, x^o) \epsilon K_D$ such that $f_{x^o} x^o = \pi^o b$, for some $f_{x^o} \epsilon F_{x^o}$.

Let $f(x^o) = \text{Min } f(x)$ for x in K. By (16), $\exists f_{x_0}$ such that

$$f_{x^o} x^o = \underset{x \epsilon K}{\text{Minimum}} \; f_{x^o} x$$

The dual of the linear program

$$\text{Minimize } f_{x^o} x \text{ subject to } Ax = b, \; x \geq 0$$

is

$$\text{Maximize } \pi b \text{ subject to } \pi A \leq f_{x^o}$$

Let π^o be the optimal solution of the maximization problem, then $\pi^o b = f_{x^o} x^o$. Moreover, it suffices to remark that $(\pi^o, x^o) \epsilon K_D$ to complete the proof.

If K is not compact, $f(x)$ may fail to attain its minimum on K. In that case, one replaces Min by Inf and adjusts the resulting proofs. This infimum may be finite or infinite. The solvability of problem (1) is answered by Proposition (16).

If $f(x)$ is differentiable, these duality statements become special cases of the results of G. B. Dantzig, E. Eisenberg, and R. Cottle [1], extended later by A. Whinston.

REFERENCES

[1] G. B. Dantzig, E. Eisenberg and R. W. Cottle: "Symmetric Dual Non-Linear Programs", Operations Research Center, University of California, RR 30 (December 1962).

[2] G. Minty: "On the Monotonicity of the Gradient of a Convex Function", Pacific Journal of Mathematics, Vol. 14, No. 1.

ACKNOWLEDGMENTS

I am very grateful to Professor G. B. Dantzig and Dr. D. W. Walkup for their valuable comments and suggestions.