NOTICE: When government or other drawings, specifications or other data are used for any purpose other than in connection with a definitely related government procurement operation, the U. S. Government thereby incurs no responsibility, nor any obligation whatsoever; and the fact that the Government may have formulated, furnished, or in any way supplied the said drawings, specifications, or other data is not to be regarded by implication or otherwise as in any manner licensing the holder or any other person or corporation, or conveying any rights or permission to manufacture, use or sell any patented invention that may in any way be related thereto.
This is a working paper. It may be expanded, modified, or withdrawn at any time. The views, conclusions, and recommendations expressed herein do not necessarily reflect the official views or policies of the United States Air Force.
A possible method for compensating for uncertainty in linear-programming problems is to replace the random elements by expected values or by pessimistic estimates of these values, or to recast the problem into a two-stage program so that, in the second stage, one can compensate for "inaccuracies" in the first stage. The purpose of RM-2751 is to examine the last of these methods in detail. More precisely, it investigates the conditions under which the first-stage decisions are optimal. In addition, formulas for using various existing computational algorithms to obtain an optimal solution are given.
ON THE SOLUTION OF TWO-STAGE LINEAR PROGRAMS UNDER UNCERTAINTY

Notes on Linear Programming and Extensions — Part 55

George B. Dantzig
Albert Madansky

RM-2751
August 10, 1961

Assigned to

This research is sponsored by the United States Air Force under contract No. AF 49(638)-700 monitored by the Directorate of Development Planning, Deputy Chief of Staff, Development, Hq USAF.

This is a working paper. It may be expanded, modified, or withdrawn at any time. The views, conclusions, and recommendations expressed herein do not necessarily reflect the official views or policies of the United States Air Force.
SUMMARY

In this paper, we study the linear programming problem

\[ Ax + By = b, \]
\[ x \geq 0, y \geq 0, \]
\[ c'x + \mathbf{E} \min_{y} f'y = \min, \]

where \( b \) is a random vector. Many short-range inventory problems can be expressed mathematically as such a program. Conditions are given for optimality of \( x \) based on the optimal dual vectors \( \pi(b, x) \) of the program

\[ By = b - Ax, \]
\[ y \geq 0, \]
\[ f'y = \min. \]

In case \( b \) has a finite discrete distribution, an explicit relationship between the optimal dual vectors for the two problems is obtained. Finally, formulas for using various existing computational algorithms, when appropriate, for obtaining an optimal \( x \) are given.
CONTENTS

SUMMARY............................................................ 111

Section
  1. INTRODUCTION.................................................. 1
  2. OPTIMALITY CONDITIONS........................................ 4
  3. COMPUTATIONAL PROCEDURES.................................... 14
  4. DISCUSSION..................................................... 22

REFERENCES......................................................... 23

LIST OF RAND NOTES ON LINEAR PROGRAMMING AND
EXTENSIONS............................................................ 25
ON THE SOLUTION OF TWO-STAGE LINEAR PROGRAMS UNDER UNCERTAINTY

1. INTRODUCTION

In this paper, we shall study the program with constraints given by

\[ \begin{align*}
    Ax + By &= b, \\
    x &\geq 0, \quad y \geq 0,
\end{align*} \]

(1)

where \( A \) and \( B \) are known \( m \times n_1 \) and \( m \times n_2 \) matrices, \( x \) and \( y \) are \( n_1 \)- and \( n_2 \)-dimensional vectors, and \( b \) is a random \( m \)-dimensional vector with known distribution. We wish to minimize, with respect to \( x \), the expression

\[ E \min_y (c'x + f'y), \]

(2)

where \( c \) and \( f \) are known \( n_1 \)- and \( n_2 \)-dimensional vectors, \( E \) denotes expectation with respect to the distribution of \( b \), and \( \theta' \) denotes the transpose of \( \theta \).

As an example of a situation leading to such a program, consider the set of possible polyhedra given by \( Ax = b, \ x \geq 0 \), where \( b \) is random and, in contrast to the usual case [6] in which one minimizes \( c'x \) subject to \( x \) lying in the intersection, over \( b \), of these polyhedra, one is instead allowed, after selecting an \( x \) and subsequently observing \( b \), to compensate with

---

1 The authors wish to acknowledge the valuable comments and suggestions of their colleagues D. R. Fulkerson and Philip Wolfe.
a vector \( y \geq 0 \) for infeasibility of the selected \( x \), at a penalty cost \( f'y \), where \( f \geq 0 \). In this case \( By \) would be \( y^+ - y^- \) and the vector \( y \) that yields the smallest penalty cost for each \( b \) and \( x \) would be composed of two parts,
\[
y^+ = b - Ax, \quad y^- = 0 \quad \text{if} \quad b \geq Ax \quad \text{or} \quad y^- = Ax - b, \quad y^+ = 0 \quad \text{if} \quad b < Ax.
\]
Since the choice of \( y \) depends on \( b \) as well as \( x \), we alter the objective from minimizing \( c'x \) to minimizing \( c'x \) plus the expected smallest penalty cost.

Many short-range inventory problems can be expressed mathematically as such a program. The vector \( x \) may represent an inventory that is to be bought at cost \( c'x \) before the random demand \( b \) is observed. Once \( b \) is observed, one must compensate by a vector \( y \) (at cost \( f'y \)) for imbalances \( (b - Ax) \) between the original inventory and the demand so as to satisfy (1). For example, coordinates of this vector \( y \) may represent the amount of additional inventory to be bought immediately to meet the excess of demand over supply or the amount of inventory to discard in case of an excess of inventory over demand.

As the structure of the problem involves a decision \( x \) to be made first, after which the random vector \( b \) is observed and a second decision \( y \) is made, we term this a two-stage problem. It may be that some coordinates of \( b \) are not random and that the corresponding equations do not involve \( y \). We shall call these equations "fixed constraints" on \( x \).

The structure of the matrices \( A \) and \( B \) may impose further constraints on \( x \). For example, \( B \) may be a positive matrix, in
which case $Ax + By = b$ and $y \geq 0$ imply that $Ax \leq b$. We shall call such constraints "induced constraints." These constraints of course may depend on the value of the random vector $b$.

If some equations do not involve $y$ but the corresponding subvector of $b$ is random, we wish to restrict our attention to the modified problem in which we take as a fixed constraint on $x$ the provision that $x$ satisfies some specific one of all these equations. We further assume that, for each $x \geq 0$ and satisfying all existing fixed and all possible induced constraints, and for each $b$, there exists a $y$ such that $(x, y)$ is feasible, i.e., satisfies (1).

This last assumption may seem very restrictive, as it says that, given any feasible $x$ and possible $b$, the set of linear equations $By = b - Ax$ must have a nonnegative solution. This assumption is motivated by the desire to solve a class of problems that can be expressed as the program given in (1) and for which this assumption is necessary for the problem to have a solution. The constraints of this class of problems have the structure

$$
\begin{align*}
A_{x1}x &= b_1, \\
A_{21}x + A_{22}y &= b_2, \\
x \geq 0, & y \geq 0,
\end{align*}
$$

where $b_1$ is a known vector and $b_2$ is a random vector with known distribution. It is clear that these constraints can be written as those of the program in (1) and that in this format, if the
objective function for (3) is of the form of (2), we have a two-stage program with fixed constraints on $x$. A further description of the underlying problem giving rise to the structure in (3) and an example of such a problem are presented in [2].

As an alternative to this assumption, we can define $K$ as the convex set of $x$ such that each $x \in K$ is nonnegative and has an associated $y$ for each $b$ such that $(x, y)$ is feasible. The problem is, then, to find $x \in K$ minimizing $c'x + E \min f'y$. These $x$ certainly satisfy all fixed and induced constraints.

2. OPTIMALITY CONDITIONS

Let us first study the program,

$$\begin{align*}
By &= b - Ax, \\
y &\geq 0, \\
f'y &= \text{min},
\end{align*}$$

(4)

and its dual for each $b$ and $x$. We assume the existence of a $\pi$ that is feasible for the dual of the program given by (4) for each $b$ and $x$. By the existence theorem for linear programming, this assumption plus the feasibility property of the convex set $K$ for the program given by (1) and (2) guarantees the existence of at least one optimal dual vector $\pi = \pi'(b, x)$ maximizing $\pi'(b - Ax)$ subject to $\pi'B \leq f'$.

By the duality theorem,

$$\begin{align*}
\min_{y} f'y &= \pi'(b, x)(b - Ax),
\end{align*}$$

(5)

so that
\[ \min_{y} c^'x + f^'y = c^'x + \pi^'(b, x)(b - Ax) = C(b, x), \]
say, and

\[ EC(b, x) = c^'x + \pi^'(b, x)(b - Ax). \]

The optimal \( x \) is then the \( x \) that minimizes \( EC(b, x) \) subject to \( x \in K \).

The following theorem establishes an immediate necessary condition for some vector \( x = \bar{x} \) to be optimal.

**Theorem 1.** Let \( \bar{x} \) be optimal and let \( \pi(b, x_1) \) be any vector that optimizes the dual of (4) for given \( b \) and \( x = x_1 \), where \( x_1 \in K \). Then

\[ [c^' - \pi^'(b, x_1)A] \bar{x} \leq [c^' - \pi^'(b, x_1)A]x_1. \]

**Proof.** Since \( \bar{x} \) is optimal and \( x_1 \) is feasible, we have

\[ c^'\bar{x} + \pi^'(b, \bar{x})(b - A\bar{x}) \leq c^'x_1 + \pi^'(b, x_1)(b - Ax_1). \]

Also, since \( \pi^'(b, \bar{x}) \) optimizes the dual of (4) when \( x = \bar{x} \), we have

\[ \pi^'(b, \bar{x})(b - A\bar{x}) \geq \pi^'(b, x_1)(b - Ax_1). \]

Hence

\[ c^'\bar{x} + \pi^'(b, x_1)(b - A\bar{x}) \leq c^'x_1 + \pi^'(b, x_1)(b - Ax_1). \]

---

1 When for a given \( b \) and \( x \) there are many optimal \( \pi \)'s, we shall mean by \( \pi(b, x) \) any vector chosen from these optima, unless some explicit statement to the contrary is made.
The following lemmas enable us to obtain another necessary condition for some vector \( x = \overline{x} \) to be optimal.

**Lemma 1.** \( EC(b, x) \) is a convex function of \( x \).

**Proof.** It is easy to check directly from the definition of convexity that \( C(b, x) \) is convex in \( x \) for each \( b \). Then \( EC(b, x) \) is a convex function of \( x \).

**Lemma 2.** The plane \( z = \langle c' - \overline{E}'(b, x_1)A \rangle x + \overline{E}'(b, x_1)b \) is a support plane to \( EC(b, x) \) at \( x = x_1 \).

**Proof.** Since at \( x = x_1 \) this plane intersects \( EC(b, x) \), all we need show is that, for \( x \neq x_1 \),

\[
(11) \quad \langle c' - \overline{E}'(b, x_1)A \rangle x + \overline{E}'(b, x_1)b \leq \langle c' - \overline{E}'(b, x)A \rangle x + \overline{E}'(b, x)b.
\]

But this is true if and only if

\[
(12) \quad \overline{E}'(b, x_1)(b - Ax) \leq \overline{E}'(b, x)(b - Ax),
\]

and this is so, since \( \overline{\pi}(b, x) \) is optimal for the dual of (4).

A consequence of these lemmas is the following theorem.

**Theorem 2.** Let \( \overline{x} \) be a relative interior point of \( K \) and let \( EC(b, x) \) be differentiable in the neighborhood of \( \overline{x} \). Then there exists a \( \overline{\pi}(b, x) \) such that \( c' - \overline{E}'(b, \overline{x})A = 0 \) if and only if \( \overline{x} \) is optimal.
Proof. Since the convex function $EC(b, x)$ is differentiable at $x$, the supporting hyperplane
\begin{equation}
 z = [c' - E\pi'(b, x)]x + E\pi'(b, x)b
\end{equation}
is tangent to $EC(b, x)$ at $x = \overline{x}$. Hence
\begin{equation}
 \frac{\partial EC(b, x)}{\partial x} \bigg|_{x=\overline{x}} = \frac{\partial z}{\partial x} \bigg|_{x=\overline{x}} = c' - E\pi'(b, x) = 0
\end{equation}
is a necessary condition for $\overline{x}$ to be optimal. Since $\overline{x}$ is a relative interior point of $K$, the condition is also sufficient.

Returning once again to the program given in (*), let $\overline{y}(b, x)$ denote the solution of this program for given $b, x$. Then $\overline{y}(b, x)$ and $\overline{y}(b, x)$ are saddle points of the function
\begin{equation}
\psi(y, \pi|x) = f'y + \pi'(b - Ax - By).
\end{equation}

Let $\phi(x, y, \pi) = c'x + E\psi(y, \pi|x)$. We then have the following results.

**Theorem 3.** Let $\overline{x}$ be optimal for the two-stage problem and $\overline{y}(b, x)$, $\overline{y}(b, x)$ be optimal for (*) and its dual. Then
\begin{equation}
\phi(\overline{x}, \overline{y}, \pi) \leq \phi(x, y, \pi) \leq \phi(x, \overline{y}, \pi).
\end{equation}

Conversely, if there exist vectors $\overline{x} \in K$ and $\overline{y}$ feasible for the dual of (*) that satisfy (16), then $\overline{x}$ is optimal for the two-stage problem and $\overline{y}$ is optimal for the dual of (*) when $x = \overline{x}$. 
Proof. Since \( \psi(\overline{y}, \overline{v} | \overline{x}) \leq \psi(\overline{y}, \overline{v} | \overline{x}) \), we have

(17) \( \phi(\overline{x}, \overline{y}, \overline{v}) = c^T \overline{x} + E\psi(\overline{y}, \overline{v} | \overline{x}) \leq c^T \overline{x} + E\psi(\overline{y}, \overline{v} | \overline{x}) = \phi(\overline{x}, \overline{y}, \overline{v}) \).

Now write \( \phi(\overline{x}, \overline{y}, \overline{v}) \) as

(18) \( \phi(\overline{x}, \overline{y}, \overline{v}) = c^T \overline{x} + E\overline{\psi}'(b, x)(b - Ax) + E[\psi' - \overline{\psi}'(b, x)B] \overline{f}(b, x) \).

Since \( \overline{f}(b, x) \) is optimal for the dual of (4), by the duality theorem the function \( \psi' - \overline{\psi}'(b, x)B \) has the property that its \( i \)-th coordinate is zero if the \( i \)-th coordinate of \( \overline{y}(b, x) \) is positive. This and the nonnegativity of \( \overline{y}(b, x) \) imply that

(19) \( \phi(\overline{x}, \overline{y}, \overline{v}) = c^T \overline{x} + E\overline{\psi}'(b, x)(b - Ax) = EC(b, x) \)

and, by optimality of \( \overline{x} \),

(20) \( \phi(\overline{x}, \overline{y}, \overline{v}) \leq \phi(\overline{x}, \overline{y}, \overline{v}) \).

To prove the converse, we must first show that \( \overline{f}(b, \overline{x}) \) is optimal for the dual of (4). Then, since \( \phi(\overline{x}, \overline{y}, \overline{v}) = EC(b, x) \), it is immediately clear that \( \overline{x} \) is optimal for the two-stage stochastic program.

Let \( \tau^*(b, \overline{x}) \) be optimal for the dual of (4). Then

(21) \( E\tau^*(b, \overline{x})(b - Ax) \geq E\overline{\psi}'(b, \overline{x})(b - Ax) \),

and also

(22) \( [\psi' - \tau^*(b, \overline{x})B] \overline{v}(b, \overline{x}) = 0. \)
But \( \bar{v}(b, x) \) satisfies \( \phi(x, \bar{y}, \pi) \leq \phi(x, \bar{y}, \bar{v}) \), or

\[
(23) \quad E[H'/(b, x)(b - Ax)] \leq E[H'/(b, x)(b - Ax)] + E[f' - H'(b, x)B] \bar{v}(b, x).
\]

If \( \bar{v}(b, x) \) is not optimal, then \( [f' - H'(b, x)B] \bar{y}(b, x) > 0 \), and hence (23) contradicts (21).

This theorem is the analogue, in the two-stage problem dealt with here, of the duality theorem for the one-stage linear team decision problem under uncertainty given in [6]. By means of this theorem, we can prove in general the following sufficient condition for optimality of \( x = \bar{x} \).

**Theorem 4.** Let \( \bar{v}(b, x) \) be optimal for the dual of (4) when \( x = \bar{x} \). If \( [e' - H'(b, x)A] \bar{x} \leq [e' - H'(b, x)A]x \) for all \( x \in K \), then \( \bar{x} \) is optimal for the two-stage problem.

**Proof.** Optimality of \( \bar{v}(b, x) \) for the dual of (4) when \( x = \bar{x} \) immediately yields the inequality

\[
(24) \quad \phi(x, \bar{y}, \pi) \leq \phi(x, \bar{y}, \bar{v}).
\]

Now by definition, hypothesis, and the facts that \( f' - H'(b, x)B \geq 0 \), \( [f' - H'(b, x)B] \bar{y}(b, x) = 0 \) for all \( x \), and \( b - Ax = E\bar{y}(b, x) = 0 \), we see that
\[ \phi(x, \bar{y}, \bar{v}) = [c' - E\pi'(b, \bar{x})A]\bar{x} + E\pi'(b, \bar{x})b \]
\[ \leq [c' - E\pi'(b, \bar{x})A]\bar{x} + E\pi'(b, \bar{x})b + E[r' - \pi'(b, \bar{x})B]y(b, x) \]
\[ (25) = c'x + E\pi'\bar{y}(b, x) + E\pi'(b, \bar{x})[b - Ax - E\bar{y}(b, x)] \]
\[ = c'x + E\pi'\bar{y}(b, x) + E\pi'(b, x)[b - Ax - E\bar{y}(b, x)] \]
\[ = \phi(x, \bar{y}, \bar{v}). \]

Hence, by Theorem 3, \( \bar{x} \) is optimal.

The foregoing result indicates that if \( \bar{x} \in K \) minimizes the linear form \( \langle c' - E\pi'(b, \bar{x})A \rangle x \), then \( \bar{x} \) is optimal. It would be worthwhile to know whether the converse of Theorem 4 is true. We have established this converse if \( E\pi'(b, x) \) is differentiable and \( \bar{x} \) exists and is an interior point of \( K \); see Theorem 2. We shall later consider the finite case. A kind of converse of Theorem 4 is the following.

**Theorem 5.** Let \( \bar{x} \) and \( \bar{y}(b, \bar{x}) \) be optimal for the two-stage problem for given \( b \), and let \( \bar{y}(b, x) \) and \( \bar{w}(b, x) \) be optimal for (4) and its dual for given \( x \). Then

\[ E[(c', f') - \pi'(b, \bar{x})(A, B)](\bar{x}, \bar{y}(b, \bar{x})) \]
\[ \leq E[(c', f') - \pi'(b, \bar{x})(A, B)](x, \bar{y}(b, x)). \]

**Proof.** It is seen, from the definition of \( \phi \) and the fact that \( b - Ax - E\bar{y}(b, x) = 0 \), that

\[ \phi(x, \bar{y}(b, x), \bar{w}(b, x)) = \phi(x, \bar{y}(b, x), \bar{w}(b, \bar{x})). \]

Now the right-hand member of (27) is the right-hand member of
(26), plus $E_x'(b, x)b$; and the left-hand member of (27) is

$\phi(x, \bar{y}, \bar{v})$, which, by Theorem 4, is greater than or equal to

$\phi(\bar{x}, \bar{y}, \bar{v})$, i.e., the left-hand member of (26) plus $E_x'(b, x)b$.

If there are only a finite number of possible $b$'s,

$b_1, \ldots, b_N$ with associated probabilities $p_1, \ldots, p_N$, where

$\sum_{i=1}^{N} p_i = 1$, then the two-stage program can be written as follows:

Find $x, y_1, \ldots, y_N$, and min $z$ satisfying

$$
\begin{align*}
Ax + By_1 &= b_1, \\
Ax + By_2 &= b_2, \\
& \vdots \\
Ax + By_N &= b_N,
\end{align*}
$$

(28)

$x \geq 0, y_1 \geq 0, y_2 \geq 0, \ldots, y_N \geq 0,

a'x + p_1f'y_1 + p_2f'y_2 + \ldots + p_Nf'y_N = z.$

As an application of the duality theorem and the optimality
test of the simplex method in this special case, we obtain the
following theorem.

**Theorem 5.** Let $\theta_1(\bar{x})$ be the $i$-th subvector ($i = 1, \ldots, N$)
in the vector of prices associated with a basic solution

$x = \bar{x}, y_i = \bar{y}_i, i = 1, \ldots, N$, for (28). Then $\bar{x}, (\bar{y}_i)$, is

optimal if

$$
\begin{align*}
c' - & \sum_{i=1}^{N} \theta_1'(\bar{x})A \geq 0, \ [c' - \sum_{i=1}^{N} \theta_1'(\bar{x})A] \bar{x} = 0, \\
p_1f' - & \theta_1'(\bar{x})B \geq 0, \ [p_1f' - \theta_1'(\bar{x})B] \bar{y}_i = 0, \ i = 1, \ldots, N.
\end{align*}
$$

(29)
Further, if \( \bar{x} \), \( (\bar{y}_1) \) is optimal, then there exist prices \( \theta_1(\bar{x}) \) satisfying (29).

We can easily prove Theorem 4 in the finite case.

**Theorem 4'.** Let \( x = \bar{x} \) and suppose that there exist optimal prices \( \bar{v}(b, \bar{x}) \) for the program in (4) such that

\[
\sigma' - \bar{v}'(b, \bar{x})A \geq 0, \quad [\sigma' - \bar{v}'(b, \bar{x})A] \bar{x} = 0.
\]

Then \( \bar{x} \) is optimal.

**Proof.** Take

\[ \theta_1(\bar{x}) = p_1 \bar{v}(b_1, \bar{x}). \]

Then

\[ \sigma - \bar{v}'(b, \bar{x})A = \sum_{i=1}^{N} p_1 \bar{v}(b_1, \bar{x}) = \sum_{i=1}^{N} \theta_1(\bar{x}). \]

Also, since \( \bar{v}'(b_1, \bar{x}) \) is optimal for the dual of (4), it satisfies

\[ f' - \bar{v}'(b_1, \bar{x})B \geq 0, \quad [f' - \bar{v}'(b_1, \bar{x})B] \bar{v}(b_1, \bar{x}) = 0. \]

Hence for \( \bar{y}_1 = \bar{v}(b_1, \bar{x}) \), the functions \( \theta_1(\bar{x}) \) so defined satisfy (29); therefore, by Theorem 5, \( \bar{x} \) is optimal.

The converse of this theorem is also easy to prove in the finite case.

**Theorem 6.** Let \( \bar{x} \) be optimal. Then there exist optimal prices \( \bar{v}(b, \bar{x}) \) for the program in (4) such that
Proof. Let \( \pi(b_1, \overline{x}) = \theta_1(\overline{x})/p_1 \), where the functions \( \theta_1(\overline{x}) \) satisfy (29). Then \( [f' - \pi^*(b_1, \overline{x})B] \overline{y}_1 = 0 \), so that \( \overline{y}_1 = \overline{y}(b_1, \overline{x}) \), and \( \pi(b_1, \overline{x}) \) is a set of optimal prices for (4).

It is interesting to contrast this necessary condition for the finite case with that of Theorem 2. In this case, not only is \( c' - E\pi'(b, \overline{x})A \geq 0 \), but further if the \( i \)-th coordinate of \( \overline{x} \) is positive then the \( i \)-th coordinate of \( c' - E\pi'(b, \overline{x})A \) is zero. In the finite case, of course, \( EC(b, x) \) does not satisfy the conditions of Theorem 2 because here \( \overline{x} \) is an extreme point of the convex region of interest and \( EC(b, x) \) is not differentiable in the neighborhood of the extreme points of this convex region.

A simple example of a two-stage problem satisfying the conditions of Theorem 2 is the problem solved in [2]. There \( x' \) was a two-dimensional nonnegative vector \( (x_1, x_2) \) with fixed constraint \( x_1 + x_2 = 100 \) so that \( K \) was a line segment in the first quadrant, and \( b \) was distributed uniformly between 70 and 80; \( EC(b, x) \) was differentiable at the optimum \( \overline{x} = (75, 25) \), a relative interior point of \( K \), and although

\[
\overline{w}(b, x) = \begin{cases} (0, \pi_2), & b \leq x_1, \\ (2, \pi_2), & b > x_1, \end{cases}
\]

was optimal for the dual of (4), where \( \pi_2 \) could take on any value, the particular \( \overline{w}(b, \overline{x}) \) for which Theorem 2 held had \( \pi_2 = 0 \).
3. COMPUTATIONAL PROCEDURES

Since the determination of an optimal \( y \), given \( x \) and \( b \), is a straightforward application of linear programming techniques to the program defined in (4), our objective is to find methods of determining an optimal \( x \) other than by solving the large program (28). As will shortly be seen, an application of the decomposition principle [4] to the dual of (28) will reduce the size of the program greatly and will directly obtain for us only the optimal \( x \) and not the optimal set of \( y \).

By virtue of Theorems 4 and 6, the problem dual to (28) can be expressed as

\[
\begin{align*}
\sum_{i=1}^{n_1} p_1^i y_1 + \sum_{i=1}^{n_2} p_2^i y_2 + \cdots + \sum_{i=1}^{N} p_N^i y_N & \leq c, \\
\sum_{i=1}^{n_1} b_1^i x_1 + \sum_{i=1}^{n_2} b_2^i x_2 + \cdots + \sum_{i=1}^{N} b_N^i x_N & \leq f, \\
\end{align*}
\]

(34)

We now note that the dual problem is in the standard form for application of the decomposition algorithm [4] to solve the program. To use the algorithm, it is convenient to call the last \( N \) sets of inequalities in (34) a single subprogram and the first set of inequalities the master program. This will reduce the program from one with \( n_1 + Nn_2 \) to one with \( n_1 + 1 \) constraints.
For notational convenience in describing the algorithm, let \( \hat{B} \) be a \( N_m \times N_n^2 \) block diagonal matrix of the form

\[
\hat{B} = \begin{bmatrix}
B & 0 \\
0 & B
\end{bmatrix}
\]

\( \hat{f} \) be a \( N_n^2 \)-dimensional vector of the form \( \hat{f} = [f' \ldots f'] \),

\( \hat{\alpha}' = [\alpha_1' \ldots \alpha_N'] \), \( \hat{\lambda}' = [p_1\lambda_1' \ldots p_n\lambda_n'] \), and \( \hat{b}' = [p_1b_1' \ldots p_Nb_N'] \).

Then (34) can be rewritten as

\[
\begin{align*}
\hat{\alpha}' \hat{\pi} & \leq c, \\
\hat{B}' \hat{\pi} & \leq \hat{f}, \\
\hat{b}' \hat{\pi} & = \max.
\end{align*}
\]

(35)

Let \( S = \{\hat{\pi} | \hat{B}' \hat{\pi} \leq \hat{f}\} \) and let \( W = \{\hat{\pi}_1, \ldots, \hat{\pi}_k\} \) be the set of extreme points of the convex set \( S \). We assume here that \( S \) is a bounded set; the slight modification in the algorithm for unbounded \( S \) is given in [4]. Also let \( \lambda_j = \langle \hat{\alpha}' \hat{\pi}_j \rangle \) and

\( r_j = \hat{b}' \hat{\pi}_j \) for \( j = 1, \ldots, k \). The extremal problem corresponding to (35) is to find numbers \( \lambda_1, \ldots, \lambda_k \) such that

\[
\begin{align*}
P_1\lambda_1 + P_2\lambda_2 + \cdots + P_k\lambda_k & \leq c, \\
\lambda_1 + \lambda_2 + \cdots + \lambda_k & = 1, \\
\lambda_1 & \geq 0, \lambda_2 \geq 0 \ldots \lambda_k \geq 0, \\
r_1\lambda_1 + r_2\lambda_2 + \cdots + r_k\lambda_k & = \max.
\end{align*}
\]

(36)

Then, as is shown in [4], \( \hat{\pi} = \sum_{j=1}^{k} \lambda_j \hat{\pi}_j \) solves (35).
Now this program has even more variables than (35), namely \( k \bar{M} \); moreover, \( k \) can be very large and hence it is not practical to determine \( W \) explicitly. Equation (36) only has \( n_1 + 1 \) constraints, however, and one need never carry more than \( n_1 + 1 \) variables in solving this equation when using the decomposition algorithm.

The algorithm is initiated once one has a feasible basis for (35), i.e., \( n_1 + 1 \) vectors \( \pi_j \) (to determine the necessary \( n_1 + 1 \) vectors \( P_j \)) and \( n_1 + 1 \) values \( \lambda_j \) that are positive and satisfy the constraints of (36). (This may be obtained using Phase One of the simplex method, for which see [4].) The prices for this basis are then determined and are used to generate those extreme points of \( S \) that appear promising and to suppress the others. This is done with each new basis formed during the process. In this case \( x \) will be the price vector for the first \( n_1 \) constraints and we shall call \( z \) the price for the last constraint. When (36) is solved, the resulting price \( x \) is optimal for the original program (28).

As the algorithm is iterative, generating new prices with each iteration, we shall append the superscript \( t \) to \( x \) and \( z \) to denote the appropriate iteration. Given the initial feasible basis, we obtain concomitantly \( x^0 \) and \( z^0 \). To test optimality of \( x^0 \) and, if not optimal, to generate a new basis, one must first find the \( \tilde{\pi} = \tilde{\pi}^0 \) that maximizes \( \tilde{\pi}' (\tilde{b} - \tilde{A} x^0) - z^0 \) subject to \( \tilde{B} \tilde{\pi} \leq \tilde{f} \). If \( \tilde{\pi}^0' (\tilde{b} - \tilde{A} x^0) - z^0 \geq 0 \), the algorithm is terminated, and \( x^0 \) solves (28). If not, a new column is added to
the $n_1 + 1$ columns of the initial basis, namely

$$
\begin{bmatrix}
\tilde{A}' \\
\pi
\end{bmatrix}
$$

and its associated cost $\tilde{b}'\pi^0$ is added to the objective function.

By means of the simplex method, a new basis and new prices $x^1, z^1$ are determined, the $\tilde{y} = \tilde{y}^1$ that maximizes $\tilde{w}'(\tilde{b} - \tilde{A}x^1) - z^1$
subject to $\tilde{B}'\tilde{y} \leq \tilde{r}$ is determined, and once again we test $\tilde{y}^1$ for optimality. The process terminates when for some iteration $t$
and prices $x^t, z^t$, we have

$$
\max_{\tilde{B}' \tilde{y} \leq \tilde{r}} \tilde{w}'(\tilde{b} - \tilde{A}x^t) + z^t \geq 0.
$$

Notice that all one needs to carry along in the computation are the columns of the current basis, $n_1 + 1$ in all.
At iteration $t$, if (37) is not satisfied, the column

$$
\begin{bmatrix}
\tilde{b}' \\
\pi
\end{bmatrix}
$$

is added to the basis and the simplex method determines the appropriate column to remove from the basis and hence from consideration in subsequent calculation.

Since

$$
\tilde{b}'\pi = \sum_{i=1}^{N} p_i \pi'(b_i, x)b_i = E\pi'(b, x)b
$$

and
(39) \[ \tilde{\pi}'Ax = \sum_{i=1}^{N} p_i \pi'(b_i, x)Ax = \tilde{E}\pi'(b, x)Ax, \]

we see that the test for optimality can be rewritten as

(40) \[ \max_{\tilde{\pi} \tilde{\pi} \in \mathbb{N}} \tilde{E}\pi'(b, x^t)b - \tilde{E}\pi'(b, x^t)Ax^t - z^t \geq 0. \]

It can also be shown that, at the optimal \( \tilde{\pi} \), we have

(41) \[ z^t = \min_{\{y_1|x^t\}} \sum_{i=1}^{N} f'y_1, \]

where \( y_1 \geq 0 \) satisfies \( B\tilde{y}_1 = b_1 - Ax^t. \)

The inequality

(42) \[ \min_{y_1 \geq 0} f'y_1 \geq \max_{B \tilde{\pi}(b_1, x) \leq f} p_i \pi'(b_i, x)(b - Ax) \]

holds for any feasible \( y_1 \) and \( \pi_1 \) of the subprogram. Summing (42) for \( i = 1, \ldots, N \) yields (40) with the inequality reversed. Hence the test for optimality can be interpreted as checking whether equality holds in (40) or, when we substitute the value of \( z^t \) given in (41), whether \( \{\pi(b_i, x^t), i = 1, \ldots, N\} \) satisfies

(43) \[ \min_{y} f'y = \tilde{E}\pi'(b, x^t)(b - Ax^t), \]

i.e., is optimal for the dual of (4), and whether

(44) \[ \tilde{E}\pi'(b, x^t)A \leq c. \]
In other words, the test for optimality of the decomposition algorithm is an implementation of the sufficient condition for optimality given in Theorem 4.

As $EC(b, x)$ is a convex function of $x$ defined over a convex set, the problem of minimizing $EC(b, x)$ is a problem in convex programming. The following procedure for solving this problem is an application of Lemma 2 and a technique due to Kelley [5]; see also [7].

We assume here that aside from the condition $x \geq 0$, there are enough fixed and induced constraints on $x$ so that $x$ lies in a bounded convex polyhedron defined by, say, $Ax = b$. The problem is then the following:

$$
\begin{align*}
\bar{A}x &= b, \\
EC(b, x) &\leq z, \\
x &\geq 0, \\
z &= \text{min}.
\end{align*}
$$

(45)

To initiate the algorithm, let $x^0$ be feasible. Consider the linear program

$$
\begin{align*}
\bar{A}x &= b, \\
\alpha 'x + E^{-1}(b - Ax) &\leq z, \\
x &\geq 0, \\
z &= \text{min}.
\end{align*}
$$

(46)

Let $x^1$ solve this program, and now consider the program
\[ \begin{align*}
\bar{A}x &= \bar{b}, \\
\begin{array}{rl}
\begin{array}{c}
c'x + E \bar{w}(b, x^0)(b - Ax) \leq z, \\
(47)
c'x + E \bar{w}(b, x^1)(b - Ax) \leq z,
\end{array}
\end{array}
\end{align*} \]

\[ x \geq 0, \]
\[ z = \text{min}. \]

Let \( x^2 \) solve this program.

One sees that on the \( k \)-th iteration of this procedure one solves the program

\[ \begin{align*}
\bar{A}x &= \bar{b}, \\
\begin{array}{rl}
\begin{array}{c}
c'x + E \bar{w}(b, x^0)(b - Ax) \leq z, \\
(48)
c'x + E \bar{w}(b, x^1)(b - Ax) \leq z,
\end{array}
\end{array}
\end{align*} \]

\[ x \geq 0, \]
\[ z = \text{min}. \]

Kelley has shown that \( \lim_{k \to \infty} \bar{w}(b, x^k) \) is the minimum of \( \bar{w}(b, x) \), though \( \lim_{k \to \infty} x^k \) does not necessarily solve the convex program.

An alternative procedure [3] rewrites (45) as

\[ \begin{align*}
\bar{A}x &= \bar{b}, \\
\begin{array}{rl}
\begin{array}{c}
x \geq 0, \\
(49)
w + \bar{w}(b, x) = 0,
\end{array}
\end{array}
\end{align*} \]

\[ w = \text{max}, \]

and uses the decomposition principle with \( \bar{A}x = \bar{b}, x \geq 0 \) as the master program and \( w + \bar{w}(b, x) = 0 \) as the single subprogram.

The extremal problem equivalent to the master program is
\[ \begin{align*}
\lambda_1 + \cdots + \lambda_k &= 1, \\
(\bar{A}x^1)\lambda_1 + \cdots + (\bar{A}x^k)\lambda_k &= \bar{b}, \\
EC(b, x^1)\lambda_1 + \cdots + EC(b, x^k)\lambda_k &= \text{min},
\end{align*} \]

where the values \( x^1, \ldots, x^k \) need not be feasible, provided some convex combination of them is. Once this program is solved and prices \( \rho^k_0, \rho^k \) are generated, one must solve the sub-program

\[ \begin{align*}
w + EC(b, x) &= 0, \\
w + \rho^k_0 &\bar{A}x + \rho^k &= \text{max},
\end{align*} \]

or, equivalently, one must find \( x = x^{k+1} \) such that

\[ \delta_k(x) = EC(b, x) - \rho^k_0 - \rho^k \]

is a minimum. If for all \( x \) we have

\[ EC(b, x) \geq \rho^k_0 - \rho^k \]

that \( EC(b, x^{k+1}) \) is minimal. If not, another column

\[ \begin{bmatrix} 1 \\ \bar{A}x^{k+1} \end{bmatrix} \]

is added to (50), \( EC(b, x^{k+1}) \lambda_{k+1} \) is added to the cost form in (50), and the algorithm is iterated. In [3] it is shown that

\[ \lim_{k \to \infty} EC(b, x^k) \]

is the minimum of \( EC(b, x) \).
4. DISCUSSION

It is well known (see, e.g., [2]) that replacing the random \( b \) by its expected value is of little help in solving the stochastic linear program. We have seen from the above discussion, however, that the expected values of the prices, with \( x \) fixed for the program given in (4), play a critical role in solving the stochastic linear program. In fact, roughly speaking, the vector

\[
(c' - ET^b_l(b, x)A')
\]

acts as the gradient of the function \( EC(b, x) \). An interesting area for future consideration is the effect of sampling from the distribution of \( b \), estimating \( ET^b_l(b, x) \) and \( ET^{b'}(b, x)b \) for each value of \( x \) generated by the iterative procedures given above, and using these estimates as \( ET^l(b, x) \) and \( ET^{b'}(b, x)b \) in those procedures.
REFERENCES


LIST OF RAND NOTES ON LINEAR PROGRAMMING
AND EXTENSIONS


RM-1266 Part 3: Computational Algorithm of the Simplex Method by G. B. Dantzig, October 26, 1953. (ASTIA No. AD 114136)


RM-1274 Part II: Composite Simplex-Dual Simplex Algorithm—I, by G. B. Dantzig, April 26, 1954. (ASTIA No. AD 114140)


Part 18: Status of Solution of Large-scale Linear Programming Problems, by G. B. Dantzig, November 30, 1954. (ASTIA No. AD 86396)


RM-1452  Part 25: The Elimination Form of the Inverse and Its Application to Linear Programming, by H. M. Markowitz, April 8, 1955. (ASTIA No. AD 86956)


RM-1736  Part 32: Solving the Transportation Problem, by L. R. Ford and D. R. Fulkerson, June 20, 1956. Published in Management Science, Vol. 3, No. 1, October, 1956, pp. 24-32. (ASTIA No. AD 111816)


Part 37: Concerning Multicommodity Networks, by J. T. Robacker, September 26, 1956. (ASTIA No. AD 112392)


Part 42: Linear Programming and Structural Design, by W. Prager, December 3, 1957. (ASTIA No. AD 150661)

RM-2158 Part 44: Transient Flows in Networks, by D. Gale, April 11, 1958. (ASTIA No. AD 150686)


RM-2597 Part 54: An Algorithm for the Mixed Integer Problem by Ralph Gomory, July 7, 1960. (ASTIA No. AD 243215)


RM-2752 Part 56: Methods of Solution of Linear Programs under Uncertainty, by Albert Madansky, April 6, 1961.