A Study on Linear Programming under Uncertainty

A Study on Linear Programming under Uncertainty

Yoshisada MUROTSU* and Fuminori OBA**

(Received June 15, 1974)

This paper presents a method to find the optimal solution to linear programming under uncertainty. A total cost is defined by adding penalty costs to activity costs when the constraints are violated, and a stochastic programming problem is set up to minimize the expected total cost. It is shown that 1) when the random coefficients have discrete distribution, the problem is reduced to the ordinary deterministic linear programming and thus the optimal solution is obtained by using the simplex method and 2) when the random coefficients have continuous distribution, the problem is proved to be a convex program and an algorithm using the gradients is presented for the case of Gaussian random variables. Further, some comment is added on the relation between the problem and the two-stage problem by Dantzig and Madansky.

1. Introduction

Many decision problems formulated as linear programming contain some parameters with uncertainty because of the measurement errors, the estimation errors, etc.[1)~4)] An approach to such a problem is to treat the uncertain parameters as random variables. The study so far made is classified into three types[5)]: 1) replacing the random variables by their particular values such as means, 2) transforming the uncertain constraints into their probability constraints and 3) recasting the problem into a two-stage problem where, in the second stage, one compensates for inaccuracies in the first stage activities. Extensive studies have been made on the problem,[3)~17] and a general survey[18)] and a monograph[19)] have been published recently.

On the other hand, somewhat different approach to the problem is proposed by Hadley,[4)] in which a total cost is defined by adding penalty costs to activity costs when the constraints are violated and a problem is set up to minimize the expected total cost. As pointed in reference 4, the problem, however, becomes nonlinear and there have not been any general techniques for finding the optimal solution. In this paper, the property of the expected total cost is investigated and the computational method is presented for the cases when the random variables have discrete or continuous distribution. It is shown that 1) when the random variables are discrete, the problem is reduced to the ordinary deterministric linear programming and thus the simplex method is applicable and 2) when the random variables are continuous, the problem is proved to be a convex program and an algorithm using the gradients is presented for the case of Gaussian random

* Department of Naval Architecture, College of Engineering.
** Department of Aeronautical Engineering, College of Engineering.

variables. Further, the relation between the present problem and the two-stage problem[8] is discussed.

2. Statement of the problem

First consider the standard linear programming problem.

PROBLEM 1: Under the constraints

$$Ax \geq b \tag{1}$$

$$x \geq 0\,, \tag{2}$$

find the control vector x to minimize the objective function

$$z = \langle c \cdot x \rangle = \sum_{i=1}^{n} c_i x_i\,, \tag{3}$$

where $\langle\ \rangle$ represents the inner product of the vectors and

$x = \text{col}\,(x_i) =$ n dimensional control vector,
$b = \text{col}\,(b_i) =$ m dimensional coefficient vector,
$c = \text{col}\,(c_i) =$ n dimensional cost coefficient vector,
$A = (a_{ij}) =$ m$\times$n coefficient matrix.

Let us now imagine that the control variables $x_i (i=1, 2, \cdots, n)$ have been determined by some means or other. However, when the coefficients are not deterministic, i.e., random variables with some distribution, the constraints may not always be satisfied for all realizations of the coefficients. Even if this can be done, the decision under these severe conditions may be too conservative to be practical. In practice, we may select the decision that will possibly result in a low expected-cost even if there may be some possibility of risk to violate the constraints. Thus we have a stochastic programming problem to determine the decision based on the trade-off between the expected cost and the expected risk. An approach to the problem is, as originally formulated by Hadley[4], to add penalty costs to the activity costs when the constraints are violated. This formulation may be justified since there may be some additional costs associated with emergency actions due to violation of the constraints. Thus we formulate the stochastic programming problem as mentioned above.

First, define m dimensional vector

$$\varepsilon = \text{col}\,(\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_m)$$

by

$$\varepsilon = b - Ax\,. \tag{4}$$

Thus $\varepsilon_i > 0$ means that the i-th constraint of (1) is not satisfied. Assume that the penalty function of the i-th constraint is given as $\psi_i(\varepsilon_i)$. Then the total cost is defined by

$$z_0 = \langle c \cdot x \rangle + \sum_{i=1}^{m} \psi_i(\varepsilon_i) I(\varepsilon_i)\,, \tag{5}$$

where

$$\mathit{1}(\varepsilon_i) = \begin{cases} 1 & \varepsilon_i > 0 \\ 0 & \varepsilon_i \leq 0\,, \end{cases} \tag{6}$$

$\psi_i(\varepsilon_i)$ = monotonously increasing convex function with sufficient smoothness, (7)

$$\psi_i(0) = 0\,. \tag{8}$$

In particular, when the penalty function is linear, we have

$$\psi_i(\varepsilon_i) = q_i \varepsilon_i\,, \tag{9}$$

where $q_i(>0)$ is the penalty coefficient.

Since the elements of A and b are random variables, those of ε are also random variables. Denoting the probability density function of ε_i by $f_i(\varepsilon_i)$, we obtain the expected value of the total cost

$$E(z_0) = \langle c \cdot x \rangle + \sum_{i=1}^{m} \int_0^\infty \psi_i(\varepsilon_i) f_i(\varepsilon_i) d\varepsilon_i\,. \tag{10}$$

In the foregoing discussion, we presume that all the constraints are random variables. In general, there are some constraints which are deterministic. Thus we divide the constraints into two sets depending on whether they are deterministic or random, i.e.,

$$\left.\begin{aligned} I_d &= \{i \mid \varepsilon_i \text{ is deterministic}\} \\ I_r &= \{i \mid \varepsilon_i \text{ is random}\} \end{aligned}\right\}, \tag{11}$$

where $\{(\cdot) \mid (\cdot\cdot)\}$ denotes the set of $(\cdot)$ that satisfies the condition $(\cdot\cdot)$. For the case, the expected total cost becomes

$$E(z_0) = \langle c \cdot x \rangle + \sum_{i \in I_r} \int_0^\infty \psi_i(\varepsilon_i) f_i(\varepsilon_i) d\varepsilon_i\,, \tag{12}$$

where $\sum_{i \in I_r}$ means that the summation is carried out over the set I_r.

Finally, the stochastic programming is formulated as follows.

PROBLEM 2: When some elements of A and b are random variables with known probability distribution, find the control variables $x_i (i=1, 2, \cdots, n)$ to minimize the expected total cost $E(z_0)$ given by Eq. (12) under the deterministic constraints

$$\left.\begin{aligned} &\varepsilon_i = b_i - \sum_{j=1}^{n} a_{ij} x_j \leq 0 \quad (i \in I_d) \\ &x_j \geq 0 \quad (j = 1, 2, \cdots, n)\,. \end{aligned}\right\}. \tag{13}$$

It should be noted here that the constraints under uncertainty in PROBLEM 1 are imbedded in the expected total cost and thus they are omitted from the constraints.

3. The case of discrete random variables

A. Optimality Condition

Consider the case where the random variables have discrete probability distribution and the penalty functions are given by (9). First we assume for brevity that only the coefficient a_{11} is a random variable with probability distribution:

$$P(a_{11}{}^{(k)}) = p_k \quad (k = 1, 2, \cdots, l), \quad \sum_{k=1}^{l} p_k = 1, \tag{14}$$

where $P(a_{11}{}^{(k)})$ is the probability that $a_{11}=a_{11}{}^{(k)}$.

Now we define the vectors:

$$\left.\begin{aligned} A_j &= \text{col}\,(a_{1j}, a_{2j}, \cdots, a_{mj}) \quad (j = 2, 3, \cdots, n) \\ A_0 &= b = \text{col}\,(b_1, b_2, \cdots, b_m) \end{aligned}\right\} \tag{15}$$

and the vector corresponding to each realization of a_{11}:

$$A_1{}^{(k)} = \text{col}\,(a_{11}{}^{(k)}, a_{21}, \cdots, a_{m1}) \quad (k = 1, 2, \cdots, l). \tag{16}$$

For each realization of a_{11}, (4) is written as follows:

$$\varepsilon^{(k)} = A_0 - \{A_1{}^{(k)}x_1 + A_2x_2 + \cdots + A_nx_n\}, \tag{17}$$

where

$$\varepsilon^{(k)} = \text{col}(\varepsilon^{(k)}, \varepsilon_2, \cdots, \varepsilon_m) \quad (k = 1, 2, \cdots, l). \tag{18}$$

The expected total cost yields

$$E(z_0) = \langle c \cdot x \rangle + q_1 \sum_{k=1}^{l} \varepsilon_1{}^{(k)} p_k \mathbf{1}(\varepsilon_1{}^{(k)}). \tag{19}$$

It must be remembered here that the constraints: $\varepsilon_2 \leq 0$, $\varepsilon_3 \leq 0$, $\cdots$, $\varepsilon_m \leq 0$ are deterministic.

Substituting (17) and (18) into (19) yields

$$\begin{aligned} E(c_0) &= \langle c \cdot x \rangle + q_1 \sum_{k=1}^{l} \{b_1 - a_{11}{}^{(k)}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n)\} p_k \mathbf{1}(\varepsilon_1{}^{(k)}) \\ &= \{c_1 - q_1 \sum_{k=1}^{l} a_{11}{}^{(k)} p_k \mathbf{1}(\varepsilon_1{}^{(k)})\} x_1 \\ &\quad + \sum_{j=2}^{n} \{c_j - q_1 a_{1j} \sum_{k=1}^{l} p_k \mathbf{1}(\varepsilon_1{}^{(k)})\} x_j + q_1 b_1 \sum_{k=1}^{l} p_k \mathbf{1}(\varepsilon_1{}^{(k)}). \end{aligned} \tag{20}$$

Since the function $\mathbf{1}(\varepsilon_1{}^{(k)})$ is 0 or 1 depending on the sign of $\varepsilon_1{}^{(k)}$, we find that the expected total cost becomes linear with respect to x_i if we divide the problem into the subproblems by the sign of $\varepsilon_1{}^{(k)}$. For example, when $\varepsilon_1{}^{(1)} > 0$ and $\varepsilon_1{}^{(k)} \leq 0$ $(k=2, 3, \cdots, l)$, the constraints and the expected total cost become as follows:

$$\left.\begin{aligned} \varepsilon_1{}^{(1)} &= b_1 - \{a_{11}{}^{(1)}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n\} > 0 \\ \varepsilon_1{}^{(2)} &= b_1 - \{a_{11}{}^{(2)}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n\} \leq 0 \\ &\cdots\cdots\cdots\cdots\cdots\cdots \\ \varepsilon_1{}^{(l)} &= b_1 - \{a_{11}{}^{(l)}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n\} \leq 0 \end{aligned}\right\}, \tag{21}$$

$$E(z_0) = \{c_1 - q_1 a_{11}{}^{(1)} p_1\} x_1 + \sum_{j=2}^{n} \{c_j - q_1 a_{1j} p_1\} x_j + q_1 b_1 p_1. \tag{22}$$

Thus the subproblems are reduced to the original problem.

PROBLEM 3: Find the control variables $x_i (i=1, 2, \cdots, n)$ to minimize the objective function $E(z_0)$ given by (22) under the constraints (21) and

$$\left.\begin{array}{ll} \varepsilon_i = b_i - \sum_{j=1}^{n} a_{ij} x_j \leq 0 & (i = 2, 3, \cdots, m) \\ x_i \geq 0 & (i = 1, 2, \cdots, n)\,. \end{array}\right\}. \tag{23}$$

The optimal solution to the problem lies on the extreme point of the region given by the constraints (21) and (23). Similarly the optimal solution to each subproblem is on the extreme point of the corresponding region determined by division of the problem depending on the sign of $\varepsilon_1^{(k)}$. Thus if we can prove the continuity of the expected total cost (19), we see that the optimal solution to PROBLEM 2 is the one that gives the minimum value to the expected total cost among the optimal solutions to the subproblems.

The continuity of $E(z_0)$ is proved in the following. Since $\varepsilon_1^{(k)}$ and $E(z_0)$ are the functions of x, we denote them by $\varepsilon_1^{(k)}(x)$ and $E(z_0)\,(x)$. In order to prove the continuity of $E(z_0)$ at any point x^0 we should only prove the fact: for an arbitrary positive number e, there exists a positive number δ such that

$$|E(z_0)(x) - E(z_0)(x^0)| < e$$

in the vicinity of x^0: $||x-x^0||<\delta$, where $||(\cdot)||$ is Euclidean norm.

$\varepsilon_1^{(k)}(x)$ can be written in the form

$$\varepsilon_1^{(k)}(x) = b_1 - \langle B_1^{(k)} \cdot x \rangle\,, \tag{24}$$

where $B_1^{(k)} = \mathrm{col}\,(a_{11}^{(k)}, a_{12}, \cdots, a_{1n})$.

Thus we have

$$\begin{aligned} & |E(z_0)(x) - E(z_0)(x^0)| \\ & = |\langle c \cdot x \rangle + q_1 \sum_{k=1}^{l} \varepsilon_1^{(k)}(x) p_k \boldsymbol{1}(\varepsilon_1^{(k)}(x)) - \langle c \cdot x^0 \rangle - q_1 \sum_{k=1}^{l} \varepsilon_1^{(k)}(x^0) p_k \boldsymbol{1}(\varepsilon_1^{(k)}(x^0))| \\ & \leq |\langle c \cdot (x - x^0) \rangle + q_1 \sum_{k=1}^{l} p_k |\varepsilon_1^{(k)}(x) - \varepsilon_1^{(k)}(x^0)| \\ & = |\langle c \cdot (x - x^0) \rangle| + q_1 \sum_{k=1}^{l} p_k |\langle B_1^{(k)} \cdot (x - x^0) \rangle| \\ & \leq \{||c|| + q_1 \sum_{k=1}^{l} p_k ||B_1^{(k)}||\} ||x - x^0||\,. \end{aligned} \tag{25}$$

Hence (25) holds if we take δ such that

$$0 < \delta < \frac{e}{\{||c|| + q_1 \sum_{k=1}^{l} p_k ||B_1^{(k)}||\}}\,. \tag{26}$$

This completes the proof of the continuity of $E(z_0)$.

B. Computational Method

Division of the problem into the subproblems can be made automatically, if we introduce the slack variables:

$$\varepsilon_1^{(k)} = y_1^{(k)+} - y_1^{(k)-}, \quad y_1^{(k)+} \geq 0, \quad y_1^{(k)-} \geq 0, \qquad (k = 1, 2, \cdots, l) \tag{27}$$

Then the constraints and the objective function are rewritten in the form

$$\varepsilon_1^{(k)}(x) = b_1 - \langle B_1^{(k)} \cdot x \rangle = y_1^{(k)+} - y_1^{(k)-} \qquad (k = 1, 2, \cdots, l), \tag{28}$$

$$E(z_0) = \langle c \cdot x \rangle + q_1 \sum_{k=1}^{l} p_k y_1^{(k)+}. \tag{29}$$

Thus the problem is reduced to the ordinary deterministic linear program.

PROBLEM 4: Find $x_i (i=1, 2, \cdots, n)$ and $y_1^{(k)+}$, $y_1^{(k)-} (k=1, 2, \cdots, n)$ to minimize the objective function (29) under the constraints (23), (27) and (28).

The foregoing discussion has been limited to the case of a single random variable. The procedure, however, can be extended to the case where more coefficients are random as shown in the following.

Let us now define the set

$$\left.\begin{aligned} J_{id} &= \{j \mid a_{ij} \text{ is deterministic, } i \in I_r\} \\ J_{ir} &= \{j \mid a_{ij} \text{ is random, } i \in I_r\} \\ &= \{j_1, j_2, \cdots, j_{n_i}\} \end{aligned}\right\}, \tag{30}$$

and introduce the slack variables:

$$y_i^{(k_i^{j_1}, k_i^{j_2}, \cdots, k_i^{j_{n_i}})+} \geq 0, \quad y_i^{(k_i^{j_1}, k_i^{j_2}, \cdots, k_i^{j_{n_i}})-} \geq 0 \tag{31}$$

such that the constraints are written as

$$b_i - \{\sum_{j \in J_{id}} a_{ij} x_j + \sum_{j \in J_r} a_{ij}^{(k_i^j)} x_j\} = y_i^{(k_i^{j_1}, k_i^{j_2}, \cdots, k_i^{j_{n_i}})+} - y_i^{(k_i^{j_1}, k_i^{j_2}, \cdots, k_i^{j_{n_i}})-} \tag{32}$$
$$(i \in I_r;\ k_i^{j_1} = 1, 2, \cdots, l_i^1;\ k_i^{j_2} = 1, 2, \cdots, l_i^2;\ \cdots;\ k_i^{j_{n_i}} = 1, 2, \cdots, l_i^{n_i}).$$

Then the objective function becomes

$$E(z_0) = \langle c \cdot x \rangle + \sum_{i \in I_r} q_i \sum_{k_i^{j_1}=1}^{l_i^1} \sum_{k_i^{j_2}=1}^{l_i^2} \cdots \sum_{k_i^{j_{n_i}}=1}^{l_i^{n_i}} y_i^{(k_i^{j_1}, k_i^{j_2}, \cdots, k_i^{j_{n_i}})+} p(k_i^{j_1}, k_i^{j_2}, \cdots, k_i^{j_{n_i}}), \tag{33}$$

where $p(k_i^{j_1}, k_i^{j_2}, \cdots, k_i^{j_{n_i}})$ is the joint probability to taken on the values:

$$a_{ij_1} = a_{ij_1}^{(k_i^{j_1})}, \cdots, a_{ij_{n_i}} = a_{ij_{n_i}}^{(k_i^{j_{n_i}})}.$$

Thus PROBLEM 2 is reduced to the ordinary deterministic linear programming problem.

PROBLEM 5: Find $x_i (i=1, 2, \cdots, n)$, $y_i^{(k_i^{j_1}, k_i^{j_2}, \cdots, k_i^{j_{n_i}})+}$ and $y_i^{(k_i^{j_1}, k_i^{j_2}, \cdots, k_i^{j_{n_i}})-} (i \in I_r;$ $k_i^{j_1} = 1, 2, \cdots, l_i^1;\ \cdots;\ k_i^{j_{n_i}} = 1, 2, \cdots, l_i^{n_i})$ to minimize the objective function (33) under the constraints (13), (31) and (32).

Thus the stochastic programming problem can be solved by using the well-known technique such as the simplex method.

C. Numerical Example

The objective function and the constraints are given as follows:

$$z = 2x_1 + x_2\,, \tag{34}$$

$$a_{11}x_1 - x_2 \geq 0\,, \quad x_1 + x_2 \geq 0\,, \quad x_1 \geq 0\,, \quad x_2 \geq 0\,, \tag{35}$$

where a_{11} is a random variable which takes the values 1 and 2 with probability: $P(a_{11}=1) = p_1$, $P(a_{11}=2) = p_2$ $(p_1 + p_2 = 1)$. The penalty cost is given as $q_1 \varepsilon_1$.

The stochastic programming problem is reduced to the following form: Find x_1 and x_2 to minimize the objective function

$$E(z_0) = \{2 - q_1 p_1 y^{(1)+} - 2q_1 p_2 y^{(2)+}\}x_1 + \{1 + q_1 p_1 y^{(1)+} + q_1 p_2 y^{(2)+}\}x_2\,, \tag{36}$$

subject to the deterministic constraints:

$$\begin{aligned} &\varepsilon_2 = 1 - x_1 - x_2 \leq 0\,, \\ &\varepsilon_1^{(1)} = y_1^{(1)+} - y_1^{(1)-} = -x_1 + x_2\,, \\ &\varepsilon_1^{(2)} = y_1^{(2)+} - y_1^{(2)-} = -2x_1 + x_2\,, \\ &x_1 \geq 0\,, \quad x_2 \geq 0\,, \quad y_1^{(1)+} \geq 0\,, \quad y_1^{(1)-} \geq 0\,, \quad y_1^{(2)+} \geq 0\,, \quad y_1^{(2)-} \geq 0\,, \end{aligned} \tag{37}$$

Thus the problem is solved by using the simplex method. The extreme points and the feasible regions are illustrated in Fig. 1. The optimal solution is dependent on the

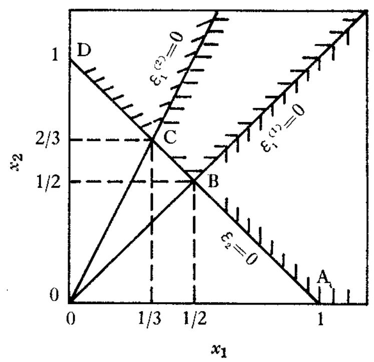

Fig. 1. Feasible region and extreme points.

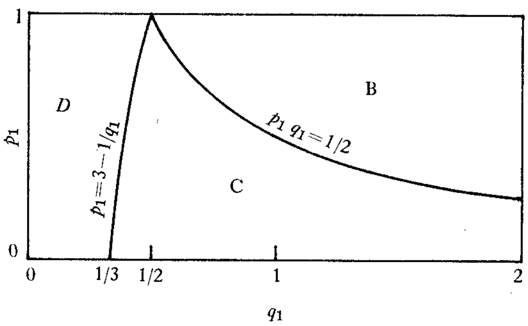

Fig. 2. Transition of optimal solution.

value of q_1 and p_1. Fig. 2 illustrates the transition of the solution. For example, when $q_1 = 5$ and $p_1 = 1/2$, the optimal solution is given as

$$x_1 = 1/2, \quad x_2 = 1/2\,, \quad y_1^{(1)+} = y_1^{(1)-} = y_1^{(2)+} = 0\,, \quad y_1^{(2)-} = 1/2\,, \tag{39}$$

and the minimum value of the expected total cost is

$$E(z_0)(1/2,\ 1/2) = 3/2\,. \tag{39}$$

4. The case of continuous random variables

For the case when the coefficients are continuous random variables, we will show that PROBLEM 2 is a convex program. Further an algorithm for solving it will be given for the case of Gaussian random variables.

A. Convexity of Expected Total Cost

In order to show that PROBLEM 2 is a convex program, it is only necessary to show that the objective function $E(z_0)$ is convex in x because the deterministic constraints are linear. The first term in $E(z_0)$ is linear, and thus we will consider the second term, i.e., the cost due to violation of the constraints. We assume that the first m_1 constraints are random and the remaining are deterministic, i.e.,

$$I_r = \{1, 2, \cdots, m_1\}, \quad I_d = \{m_1+1, \ m_1+2, \cdots, m\}. \tag{40}$$

Corresponding to the random elements of ε i.e., the first m_1 elements of ε which is denoted by ${}^r\varepsilon$, we define the matrix ${}^rA$ and the vector ${}^rb$ by

$${}^r\varepsilon = {}^rb - {}^rAx,$$

and $2m_1$ dimensional vector $\tilde{\varepsilon}$ by

$$\tilde{\varepsilon} = \text{col}\,(\varepsilon_1^+, \varepsilon_2^+, \cdots, \varepsilon_m^+, \varepsilon_1^-, \varepsilon_2^-, \cdots, \varepsilon_m^-) \tag{41}$$

where

$$\varepsilon_i^+ = \begin{cases} \varepsilon_i & \varepsilon_i > 0 \\ 0 & \varepsilon_i \leq 0, \end{cases} \qquad \varepsilon_i^- = \begin{cases} 0 & \varepsilon_i > 0 \\ -\varepsilon_i & \varepsilon_i \leq 0. \end{cases} \tag{42}$$

The penalty cost corresponding to a realization of ${}^rA$ and ${}^rb$ is denoted by $z_p(x, {}^rA, {}^rb)$, i.e.,

$$z_p(x, {}^rA, {}^rb) \overset{\Delta}{=} \sum_{i=1}^{m_1} \psi_i(\varepsilon_i) \mathbf{1}(\varepsilon_i) = \sum_{i=1}^{m_1} \psi_i(\varepsilon_i^+). \tag{43}$$

Particularly, when the penalty cost is given by (9), it becomes

$$z_p(x, {}^rA, {}^rb) = \langle q \cdot \tilde{\varepsilon} \rangle = \sum_{i=1}^{m_1} q_i \varepsilon_i^+, \tag{44}$$

where

$$q = \text{col}\,(q_1, q_2, \cdots, q_{m_1}, \overbrace{0, 0, \cdots, 0}^{m_1}). \tag{45}$$

To prove the convexity of $z_p(x, {}^rA, {}^rb)$ for an arbitrary realization of the random elements of ${}^rA$ and ${}^rb$, we will consider the following problem.

PROBLEM 6: Given a realization of ${}^rA$ and ${}^rb$, find y to minimize the objective function

$$\psi(y) = \sum_{i=1}^{m_1} \psi_i(y_i) \tag{46}$$

under the constraints

$$By = {}^r b - {}^r A x\,, \quad y \geq 0\,, \tag{47}$$

where

$$y = \mathrm{col}\,(y_1, y_2, \cdots, y_{2m_1})\,, \tag{48}$$

$$B = \left(\overbrace{\begin{matrix} 1 & & 0 \\ & 1 & \\ 0 & & \ddots \\ & & 1 \end{matrix}}^{m_1} \vdots \overbrace{\begin{matrix} -1 & & 0 \\ & -1 & \\ 0 & & \ddots \\ & & -1 \end{matrix}}^{m_1}\right) \Bigg\} m_1\,. \tag{49}$$

Since

$$\{t \mid t = \mathrm{B}y,\ y \geq 0\} = R^{m_1}\,, \tag{50}$$

PROBLEM 6 has a solution for all realization of ${}^r A$ and ${}^r b$. The optimal solution is given by

$$y = \tilde{\varepsilon} \tag{51}$$

as will be shown in the following.

From (4) and (42), the elements of $\tilde{\varepsilon}$ are expressed in the following form for any y which satisfies (47):

$$\left.\begin{aligned} \varepsilon_i^+ &= \max\{0,\ y_i - y_{m_1+i}\} \\ \varepsilon_i^- &= \max\{0,\ y_{m_1+i} - y_i\} \end{aligned}\right\} \quad (i = 1, 2, \cdots, m_1)\,, \tag{52}$$

where max $\{(\cdot), (\cdot\cdot)\}$ means the maximum value of $(\cdot)$ and $(\cdot\cdot)$. By assumption (7), $\psi_i(\varepsilon_i)$ is a monotonously increasing function of ε_i and thus the following inequality holds

$$\begin{aligned} \psi(\tilde{\varepsilon}) &= \sum_{i=1}^{m_1} \psi_i(\varepsilon_i^+) \\ &= \sum_{i=1}^{m_1} \psi_i(\max\{0,\ y_i - y_{m_1+i}\}) \\ &\leq \sum_{i=1}^{m_1} \psi_i(y_i) = \psi(y)\,. \end{aligned} \tag{53}$$

Equality holds only when $y = \tilde{\varepsilon}$. Hence the optimal solution to PROBLEM 6 is given by (51).

The objective function for the case $y = \tilde{\varepsilon}$ is given as

$$\psi(\tilde{\varepsilon}) = \sum_{i=1}^{m_1} \psi_i(\varepsilon_i^+)\,, \tag{54}$$

which, from (47), is also equal to

$$\psi(\tilde{\varepsilon}) = z_p(x, {}^r A, {}^r b)\,. \tag{55}$$

Finally we conclude that the penalty cost term in the total cost (12) is equal to the optimal value of the objective function of PROBLEM 6.

Next we will show that $z_p(x, {}^r A, {}^r b)$ is a convex function of x. Let y^1 and y^2 be the optimal solutions to PROBLEM 6 when $x = x^1$ and $x = x^2$, respectively. Although y^λ given by

$$y^\lambda = \lambda y^1 + (1-\lambda) y^2 \qquad (0 < \lambda < 1) \tag{56}$$

is a feasible solution, it is not always the optimal solution to the problem when

$$x^\lambda = \lambda x^1 + (1-\lambda)x^2 . \tag{57}$$

To prove this, we consider the sets

$$\begin{aligned} I^{1+} &= \{i \mid \varepsilon_i^{1+} > 0,\ y^1 = \bar{\varepsilon}^1,\ i \leq m_1\} \\ I^{2+} &= \{i \mid \varepsilon_i^{2+} > 0,\ y^2 = \bar{\varepsilon}^2,\ i \leq m_1\} \end{aligned} \tag{58}$$

In general, we have

$$I^{1+} \neq I^{2+} . \tag{59}$$

Hence for i which satisfies

$$i \in \{I^{1+} \cup I^{2+} - I^{1+} \cap I^{2+}\} , \tag{60}$$

ε_i^{1+}, ε_i^{2-} ($i \leq m_1$) are positive while ε_i^{2+}, ε_i^{1-} are zero and vice versa. Hence we have

$$\left.\begin{aligned} y_i^\lambda &= \lambda\varepsilon_i^{1+} + (1-\lambda)\varepsilon_i^{2+} > 0 \\ y_{m+i}^\lambda &= \lambda\varepsilon_i^{1-} + (1-\lambda)\varepsilon_i^{2-} > 0 \end{aligned}\right\} (i \leq m_1) . \tag{61}$$

On the other hand, the optimal solution to PROBLEM 6 when $x = x^\lambda$ is given in the form:

$$y = \bar{\varepsilon}^\lambda = \mathrm{col}\,(\varepsilon_1^{\lambda+}, \varepsilon_2^{\lambda+}, \cdots, \varepsilon_{m_1}^{\lambda+}, \varepsilon_1^{\lambda-}, \varepsilon_2^{\lambda-}, \cdots, \varepsilon_{m_1}^{\lambda-}) , \tag{62}$$

where

$$\left.\begin{aligned} \varepsilon_i^{\lambda+} &= \max\,\{0, y_i - y_{m_1+i}\} \\ \varepsilon_i^{\lambda-} &= \max\,\{0, y_{m_1+i} - y_i\} \end{aligned}\right\} (i = 1, 2, \cdots, m_1) . \tag{63}$$

Then the inequalities:

$$\varepsilon_i^{\lambda+} > 0 , \quad \varepsilon_i^{\lambda-} > 0 \qquad (i = 1, 2, \cdots, m_1) . \tag{64}$$

do not hold simultaneously and thus the elements of y^λ and $\bar{\varepsilon}^\lambda$ do not coincide with the elements corresponding to i given by (60). Therefore, y^λ is not always the optimal solution to PROBLEM 6 when $x = x^\lambda$.

From the foregoing discussion, we have the following inequality:

$$\begin{aligned} z_p(x^\lambda, {}^rA, {}^rb) &= \psi(\bar{\varepsilon}^\lambda) \\ &\leqq \psi(y^\lambda) \\ &\leq \lambda\psi(y^1) + (1-\lambda)\psi(y^2) \\ &= \lambda z_p(x^1, {}^rA, {}^rb) + (1-\lambda) z_p(x^2, {}^rA, {}^rb) , \end{aligned} \tag{65}$$

where the last inequality follows from the convexity of $\psi(\varepsilon)$. (65) proves that $z_p(x, {}^rA, {}^rb)$ is a convex function of x. Since the inequality (65) holds for any realizations of ${}^rA$ and ${}^rb$, the following inequality holds from the property of integral

$$E(z_p(x^\lambda, {}^rA, {}^rb)) \leq \lambda E(z_p(x^1, {}^rA, {}^rb)) + (1-\lambda) E(z_p(x^2, {}^rA, {}^rb)) , \tag{66}$$

where $E(\cdot)$ means the expectation with respect to ${}^rA$ and ${}^rb$. Thus we conclude that $E(z_p)$ is a convex function of x.

It is easily seen that PROBLEM 6 is equivalent to the second stage problem of the two-stage problem proposed by Dantzig and Madansky[8]:

PROBLEM 7: Under the constraints

$$By = {}^{r}b - {}^{r}Ax, \ \ x \geq 0, \ \ y \geq 0 \, , \tag{67}$$

and (14), minimize the objective function

$$E(\langle c \cdot x \rangle + \min_{y} \psi(y)) \, , \tag{68}$$

where $\psi(y)$ and B are given by (46) and (49), respectively.

Thus we can conclude that PROBLEM 2 is equivalent to the two-stage problem formulated as PROBLEM 7. Therefore, given the penalty function $\psi(y)$ by Eq. (7), problem 2 with continuous random variables is reduced to the special case of simple recourse[13].

B. Computational Method for Case of Gaussian Random Variables

In the previous section, it is shown that PROBLEM 2 is a convex program and thus the optimal solution can be obtained by the well-established extremum seeking methods. In the following an algorithm using the gradients is shown to be applicable to the case where a_{ij} and $b_i (i \in I_r, j \in J_{ir})$ are Gaussian random variables.

First calculate the mean and the standard deviation of ε_i:

$$\bar{\varepsilon}_i = \bar{b}_i - \sum_{j=1}^{n} \bar{a}_{ij} x_j \qquad (i \in I_r) \, , \tag{69}$$

$$\begin{aligned} \sigma_{\varepsilon_j}{}^2 = & \sum_{j=1}^{n} \sum_{k=1}^{n} \rho_{a_{ij} a_{ik}} \sigma_{a_{ik}} \sigma_{a_{ik}} x_j x_k \\ & - 2 \sum_{j=1}^{n} \rho_{a_{ij} b_i} \sigma_{a_{ij}} \sigma_{b_i} x_j + \sigma_{b_i}{}^2 \qquad (i \in I_r) \, , \end{aligned} \tag{70}$$

where $\overline{(\cdot)}$=mean of $(\cdot)$, $\sigma_{(\cdot)}$=standard deviation of $(\cdot)$

$\rho_{(\cdot)(\cdot\cdot)}$=correlation coefficient between $(\cdot)$and $(\cdot\cdot)$.

When a_{ij} and $b_i \, (i \in I_r, j \in J_{ir})$ are Gaussian random variables, ε_i is also a Gaussian random variable and the probability density function is given as

$$f_i(\varepsilon_i) = \frac{1}{\sqrt{2\pi}\, \sigma_{\varepsilon_i}} \exp \left\{ - \frac{(\varepsilon_i - \bar{\varepsilon}_i)^2}{2 \sigma_{\varepsilon_i}{}^2} \right\} . \tag{71}$$

By standardizing ε_i by the transformation:

$$u_i = (\varepsilon_i - \bar{\varepsilon}_i) / \sigma_{\varepsilon_i} \, , \tag{72}$$

we express the objective function of PROBLEM 2 as

$$E(z_0)(x) = \langle c \cdot x \rangle + \sum_{i=1}^{m_1} \int_{-\bar{\varepsilon}_i / \sigma_{\varepsilon_i}}^{\infty} \psi_i(\sigma_{\varepsilon_i} u_i + \bar{\varepsilon}_i) \phi(u_i) du_i \, , \tag{73}$$

where

$$\phi(t) = \frac{1}{\sqrt{2\pi}} \exp(-t^2/2) \, . \tag{74}$$

Thus the first derivative of $E(z_0)(x)$ is calculated as

$$\frac{\partial E(z_0)}{\partial x_j} = c_j - \sum_{i=1}^{m_1} \psi_i(0)\phi(-\bar{\varepsilon}_i/\sigma_{\varepsilon_i})\left(\frac{\bar{\varepsilon}_i}{\sigma_{\varepsilon_i}^2}\frac{\partial \sigma_{\varepsilon_i}}{\partial x_j} - \frac{1}{\sigma_{\varepsilon_i}}\frac{\partial \bar{\varepsilon}_i}{\partial x_j}\right)$$
$$+\sum_{i=1}^{m_1}\int_{-\bar{\varepsilon}_i/\sigma_{\varepsilon_i}}^{\infty} \psi_i'(\sigma_{\varepsilon_i}u_i+\bar{\varepsilon}_i)\left(u_i\frac{\partial \sigma_{\varepsilon_i}}{\partial x_j}+\frac{\partial \bar{\varepsilon}_i}{\partial x_j}\right)\phi(u_i)du_i$$
$$= c_j+\sum_{i=1}^{m_1}\int_{-\bar{\varepsilon}_i/\sigma_{\varepsilon_i}}^{\infty} \psi_i'(\sigma_{\varepsilon_i}u_i+\bar{\varepsilon}_i)\left(u_i\frac{\partial \sigma_{\varepsilon_i}}{\partial x_j}+\frac{\partial \bar{\varepsilon}_i}{\partial x_j}\right)\phi(u_i)du_i\,, \quad (75)$$

where

$$\psi_i'(\varepsilon_i) = \frac{d\psi_i(\varepsilon_i)}{d\varepsilon_i}\,, \quad (76)$$

and the last equality follows from (8). Further, we have the following relations from (68) and (69):

$$\frac{\partial \bar{\varepsilon}_i}{\partial x_j} = -\bar{a}_{ij}\,, \quad (77)$$

$$\frac{\partial \sigma_{\varepsilon_i}}{\partial x_j} = \{\sum_{k=1}^{n} \rho_{a_{ij}a_{ik}}\sigma_{a_{ij}}\sigma_{a_{ik}}x_k - \rho_{a_{ij}b_i}\sigma_{a_{ij}}\sigma_{b_i}\}\,. \quad (78)$$

When ψ_i' is differentiable, the Hessian of the objective function is calculated and the second derivatives can be obtained. Thus the various extremum seeking techniques using the gradients can be applied to the problem.

C. Numerical Example

The constraints and the objective function are given as follows:

$$\left.\begin{array}{ll} a_{11}x_1+a_{12}x_2 \geq b_1\,, & a_{21}x_1+a_{22}x_2 \geq b_2 \\ x_1 \geq 0\,, \quad x_2 \geq 0 & \end{array}\right\}, \quad (79)$$

$$z = c_1x_1+c_2x_2 = 2x_1+x_2\,, \quad (80)$$

where a_{ij} and b_i $(i, j=1, 2)$ are independent Gaussian random variables given by

$$\begin{array}{llllll} \bar{a}_{11} = 1, & \sigma_{a11} = 0.1, & \bar{a}_{12} = 1, & \sigma_{a12} = 0.1, & \bar{a}_{21} = 1, & \sigma_{a21} = 0.1 \\ \bar{a}_{22} = -1, & \sigma_{a22} = 0.1, & \bar{b}_1 = 1, & \sigma_{b_1} = 0.1, & \bar{b}_2 = 0, & \sigma_{b2} = 0.1\,. \end{array} \quad (81)$$

The mean and the standard deviation of ε_i are calculated as

$$\bar{\varepsilon}_i = \bar{b}_i - \bar{a}_{i1}x_1 - \bar{a}_{i2}x_2 \qquad (i=1, 2)\,, \quad (82)$$

$$\sigma_{\varepsilon_i}^2 = \sigma_{a_{i1}}^2x_1^2+\sigma_{a_{i2}}^2x_2^2+\sigma_{b_i}^2 \qquad (i=1, 2)\,. \quad (83)$$

When the penalty function is given by

$$\psi_1(\varepsilon_1^+) = q_1\varepsilon_1^+\,, \quad \psi_2(\varepsilon_2^+) = q_2\varepsilon_2^+\,, \quad (84)$$

the expected total cost becomes

$$E(z_0) = \sum_{j=1}^{2} c_j x_j + \sum_{i=1}^{2} q_i \int_{-\bar{\varepsilon}_i/\sigma_{\varepsilon_i}}^{\infty} (\sigma_{\varepsilon_i} u_i + \bar{\varepsilon}_i)\phi(u_i)du_i$$
$$= \sum_{j=1}^{2} c_j x_j + \sum_{i=1}^{2} q_i \left\{ \frac{\sigma_{\varepsilon_i}}{\sqrt{2\pi}} \exp\left(-\frac{\bar{\varepsilon}_i^2}{2\sigma_{\varepsilon_i}^2}\right) + \bar{\varepsilon}_i \Phi\left(\frac{\bar{\varepsilon}_i}{\sigma_{\varepsilon_i}}\right) \right\}, \tag{85}$$

where

$$\Phi(t) = \int_{-\infty}^{t} \phi(u)du\,. \tag{86}$$

The first derivative of $E(z_0)(x)$ is

$$\frac{\partial E(z_0)}{\partial x_j} = c_j + \sum_{i=1}^{2} q_i \left\{ \frac{\sigma_{a_{ij}}^2}{\sqrt{2\pi}\sigma_{\varepsilon_i}} x_j \exp\left(-\frac{\bar{\varepsilon}_i^2}{2\sigma_{\varepsilon_i}^2}\right) - \bar{a}_{ij} \Phi\left(\frac{\bar{\varepsilon}_i}{\sigma_{\varepsilon_i}}\right) \right\}. \tag{87}$$

Table 1 Optimal solution

No.	Penalty		Optimal solution		Probability		Expected
	q_1	q_2	x_1	x_2	Prob $[\varepsilon_1 \leq 0]$	Prob $[\varepsilon_2 \leq 0]$	total cost
1	5	5	0.608	0.450	0.678	0.896	1.828
2	10	10	0.667	0.459	0.835	0.947	1.933
3	100	100	0.818	0.471	0.982	0.994	2.221
4	1000	1000	0.945	0.476	0.998	0.999	2.472
5	5	10	0.631	0.427	0.676	0.948	1.849
6	5	100	0.690	0.367	0.672	0.995	1.905
7	5	1000	0.737	0.319	0.669	0.999	1.952
8	10	5	0.643	0.482	0.835	0.896	1.912
9	100	5	0.728	0.559	0.983	0.893	2.134
10	1000	5	0.794	0.618	0.998	0.892	2.318

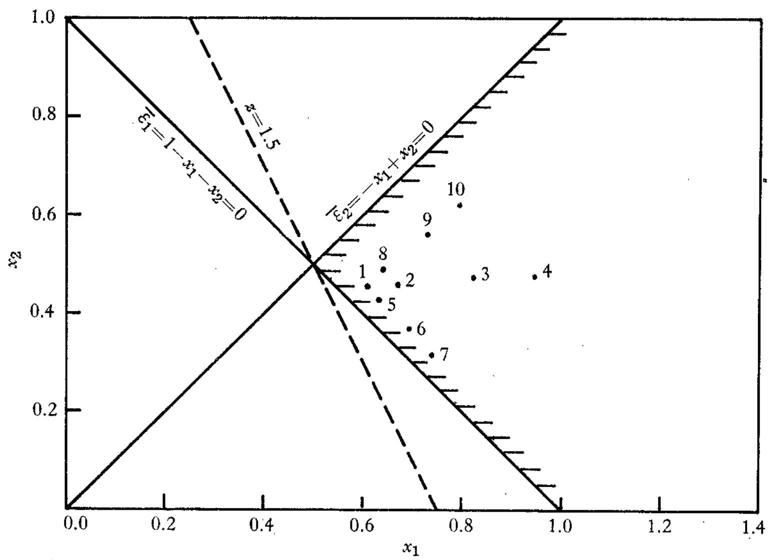

Fig. 3. Optimal solution.

The optimal solutions for various values of q_1 and q_2 are calculated by using SUMT (Sequentially Unconstrained Minimization Technique)[20]. The results are shown in Table 1 and Fig. 3. The numbers in the figure corresponds to those in the column No. in Table 1. We see that the optimal solution tends toward the safety side and thus the probability to satisfy the constraints becomes higher as the penalty becomes large.

5. Conclusion

An approach to stochastic programming is presented: A total cost is defined by adding penalty costs when the constraints are violated and a problem is set up to minimize the expected total cost. It is shown that 1) when the coefficients under uncertainty are discrete random variables, the problem is reduced to the ordinary deterministic linear program and 2) when random coefficients have continuous probability distribution, the problem is proved to be a convex program. Further, computational methods are presented for both cases.

References

1) K. Iwata, Y. Murotsu, T. Iwatsubo and S. Fujii, Trans. ASME, J. Eng. for Industry, **94**, Series B, 4, pp. 1099–1107, Nov. (1972).
2) S.E. Elmaghraby, Management Science, **6**, pp. 270–294 (1960).
3) G.B. Dantzig, Linear Programming and Extensions, Princeton, N.J.: Princeton University Press (1963).
4) G. Hadley, Nonlinear and Dynamic Programming, Reading, Mass. Addison Wesley Publishing Co. (1964).
5) A. Madansky, Operations Research, **10**, 4, pp. 463–471 (1962).
6) A. Charnes and W.W. Cooper, Management Science. **5**, 1, pp. 73–79 (1959).
7) A. Madansky, Management Science, **6**, 2, pp. 197–204 (1960).
8) G.B. Dantzig and A. Madansky, Proc. of IVth Berkeley Symp. on Mathematical Statistics and Probability, pp. 165–176 (1961).
9) A.C. Williams, J. Soc. Indust. Appl. Math., **13**, 4, pp. 927–940 (1965).
10) R.J.B. Wets, J. SIAM Appl. Math., **14**, 1, pp. 89–105 (1966).
11) A.C. Williams, J. SIAM Appl. Math., **14**, 4, pp. 668–677 (1966).
12) R.J.B. Wets, J. SIAM Appl. Math. **14**, 5, pp. 1143–1151 (1966).
13) D.W. Walkup and R.J.B. Wets, J. SIAM Appl. Math., **15**, pp. 1299–1314 (1967).
14) G.H. Symonds, Operations Research, **16**, 6, pp. 1152–1159 (1968).
15) D.W. Walkup and R.J.B. Wets, J. SIAM Appl. Math., **17**, 1, pp. 98–103 (1969).
16) D.W. Walkup and R.J.B. Wets, J. SIAM Control, **7**, 3, pp. 430–436 (1969).
17) Y. Murotsu, F. Ohba and H. Itoh, Bulletin of University of Osaka Prefecture, Series A, **20**, 1, pp. 97–108 (1971).
18) V.V. Kolbin, "Stochastic Programming" in Progress in Mathematics edited by R.V. Gamkrelidze Vol. 11, Plenum Press, New York, (1971).
19) S. Vajda, Probabilistic Programming, Academic Press, New York, (1972).
20) J. Kowalik and M.R. Osborne, Methods for Unconstrained Optimization Problems, Elsevier, Switzerland (1968).