INEQUALITIES FOR STOCHASTIC LINEAR PROGRAMMING PROBLEMS*†

ALBERT MADANSKY

The RAND Corporation

Consider a linear-programming problem in which the "right-hand side" is a random vector whose expected value is known and where the expected value of the objective function is to be minimized. An approximate solution is often found by replacing the "right-hand side" by its expected value and solving the resulting linear programming problem. In this paper conditions are given for the equality of the expected value of the objective function for the optimal solution and the value of the objective function for the approximate solution; bounds on these values are also given. In addition, the relation between this problem and a related problem, where one makes an observation on the "right-hand side" and solves the (nonstochastic) linear programming problem based on this observation, is discussed.

In [3] Dantzig presents a two-stage linear programming problem where the "right-hand side" is a random vector and the expected value of the objective function is to be minimized. It was stated there that one cannot in general replace the "right-hand side" by its expected value and obtain an optimal solution. For example, in [5] the expected value of the objective function for the optimal solution is \$1,524,000, whereas the value of the objective function when the "right-hand side" is replaced by its expected value is \$1,000,000. What we shall show is that the direction of the inequality in that example is the case in general.

This procedure of replacing the "right-hand side" by its expected value and solving this linear programming problem is quite common, however, and, as Dantzig conjectures [3], probably provides an excellent starting solution for any improvement techniques which might be devised. It is therefore of interest to examine under what conditions the expected value of the objective function for the optimal solution is close to the value of the objective function for this "starting solution." This we shall do. More precisely, we shall give conditions for the equality of these two values and also bounds for each of these values.

However, let us first consider the usual one-stage non-stochastic linear programming problem, which, for reasons that will become evident later, we write in the form:

$$Ax + By = b, \qquad x \geqq 0, \qquad y \geqq 0,$$

$$\text{minimize } c'x + f'y \text{ with respect to } x \text{ and } y$$

where A and B are known $m \times n_1$ and $m \times n_2$ matrices, respectively, b, c, and

* Received February 1959.

† I am grateful to George B. Dantzig and Philip Wolfe for their very helpful comments and suggestions.

f are known m, n_1, and n_2 dimensional vectors, and x and y are, respectively, n_1 and n_2 dimensional vectors of activity levels to be determined.

Let us denote $\min_y (c'x + f'y)$ for given x subject to the above restraints as $C(b, x)$. We shall now give a result analogous to Theorem 3 of Beale [1].

Lemma 1: $\min_x C(b, x)$ *is a convex function of* b.

Proof: Fix $b = b_1$ and define $\bar{x}(b_1)$ as the vector which minimizes $C(b, x)$ subject to $\bar{x}(b_1) \geqq 0$, $A\bar{x}(b_1) = b_1$.[1] Fix $b = b_2$ and define $\bar{x}(b_2)$ analogously. Now let $b = \lambda b_1 + (1 - \lambda)b_2$, $0 \leqq \lambda \leqq 1$. There is a vector $x(b)$ such that $x(b) = \lambda\bar{x}(b_1) + (1 - \lambda)\bar{x}(b_2) \geqq 0$ and $b = Ax(b) = \lambda A\bar{x}(b_1) + (1 - \lambda)A\bar{x}(b_2) = \lambda b_1 + (1 - \lambda)b_2$. Therefore $C(b, x(b)) = \lambda C(b_1, \bar{x}(b_1)) + (1 - \lambda)C(b_2, \bar{x}(b_2))$. But for $b = \lambda b_1 + (1 - \lambda)b_2$ there is an $\bar{x}(b)$ such that $\min_x C(b, x) = C(b, \bar{x}(b)) \leqq C(b, x(b))$, i.e.,

$$C(\lambda b_1 + (1 - \lambda)b_2, \bar{x}(b)) \leqq \lambda C(b_1, \bar{x}(b_1)) + (1 - \lambda)C(b_2, \bar{x}(b_2)) = \lambda \min_x C(b_1, x) + (1 - \lambda) \min_x C(b_2, x).$$

An alternative proof is the following. Consider the dual problem, maximize $b'z$ subject to $z'A \leqq c$, $z'B \leqq f$. If $\bar{z}$ maximizes $b'z$ and satisfies these restraints, then $b'\bar{z} = C(b, \bar{x}(b))$. When $b' = \lambda b'_1 + (1 - \lambda)b'_2$, $b'\bar{z} = \lambda b'_1\bar{z} + (1 - \lambda)b'_2\bar{z}$. If $\bar{z}_1$ maximizes b'_1z and $\bar{z}_2$ maximizes b'_2z, then

$$\min_x C(b, x) = C(b, \bar{x}(b)) = b'\bar{z} \leqq \lambda b'_1\bar{z}_1 + (1 - \lambda)b'_2\bar{z}_2 = \lambda C(b_1, \bar{x}(b_1)) + (1 - \lambda)C(b_2, \bar{x}(b_2)),$$

as above.

Lemma 2: $\min_x C(b, x)$ *is a continuous function of* b.

Proof: If $x_1 \geqq 0$ satisfies $Ax_1 = b$, and $x_2 \geqq 0$ satisfies $Ax_2 = b_2$, then $x = \lambda x_1 + (1 - \lambda)x_2$, $0 \leqq \lambda \leqq 1$, satisfies $Ax = \lambda b_1 + (1 - \lambda)b_2$. Hence the set of b's for which there exists an x which minimizes $C(b, x)$ is convex. Since $\min_x C(b, x)$ is a convex function over a convex set of b's, it is continuous at every interior point of this convex set.

Let b be on the boundary of this convex set. Consider the sequence $\{b_i\}$ such that $\lim_{i\to\infty} b_i = b$, $\bar{x}(b_i) \geqq 0$, and $A\bar{x}(b_i) = b_i$ for each i, where $\bar{x}(b_i)$ minimizes $C(b_i, x)$. Now

$$\lim_{i\to\infty} \min_x C(b_i, x) = \lim_{i\to\infty} C(b_i, \bar{x}(b_i)) \geqq \min_x C(\lim_{i\to\infty} b_i, x) = C(b, \bar{x}(b)) = \min_x C(b, x)$$

by definition of $\bar{x}(b)$.

Also by convexity of $\min_x C(b, x)$,

$$\lim_{i\to\infty} \min_x C(b_i, x) \leqq \min_x C(b, x).$$

Hence

$$\lim_{i\to\infty} \min_x C(b_i, x) = \min_x C(\lim_{i\to\infty} b_i, x).$$

[1] Without loss of generality we can disregard the By term in the proofs of Lemmas 1 and 2.

In Dantzig's two-stage stochastic linear programming problem referred to above, one must first make a decision x, then observe the random "right-hand side" b, and finally, in the second stage, compensate with activity vector y for "inaccuracies" in the first decision. The problem here is to find x subject to $Ax + By = b$, $x \geqq 0$, $y \geqq 0$, which minimizes $EC(b, x) = E \min_y (c'x + f'y)$, under the assumption that for each (b, x) there exists a y such that (x, y) is feasible.[2]

A related problem, also called a stochastic linear programming problem (cf., e.g., [11]) is the following. The linear programmer waits till an observation on the random b is made and then solves the (nonstochastic) linear programming problem based on this observed b. We shall call this the "wait-and-see" situation and, in contrast to this situation, we call the situation Dantzig treats a "here-and-now" situation.

Let $\bar{x}(Eb)$ be the decision which minimizes $C(Eb, x)$, i.e., $C(Eb, \bar{x}(Eb)) = \min_x C(Eb, x)$. Then it is obvious that

$$EC(b, \bar{x}(Eb)) \geqq \min_x EC(b, x).$$

In the example of [5], $EC(b, \bar{x}(Eb)) =$ \$1,666,000 and $\min_x EC(b, x) =$ \$1,524,000.

We also note immediately that the value of the objective function in Dantzig's two-stage "here-and-now" problem is at least as great as the expected value of the objective function of the "wait-and-see" problem, i.e.,

$$\min_x EC(b, x) \geqq E \min_x C(b, x).$$

For let $\bar{x}$ be that x which minimizes $EC(b, x)$ and let $\bar{x}(b)$ be that x which minimizes $C(b, x)$. Then $\min_x EC(b, x) = EC(b, \bar{x})$, $E \min_x C(b, x) = EC(b, \bar{x}(b))$, and, since $C(b, \bar{x}) \geqq C(b, \bar{x}(b))$ for every b, $EC(b, \bar{x}) \geqq EC(b, \bar{x}(b))$.

Finally, since by Lemmas 1 and 2 we have established that $\min_x C(b, x)$ is a continuous convex function over the convex set of b's, when b is a random vector we see by Jensen's inequality[3] that

$$E \min_x C(b, x) \geqq \min_x C(E\, b, x).$$

A direct proof of this fact due to Vajda [12], not relying on convexity and Jensen's inequality, is provided by noticing that for any b, $\min_x C(b, x) = b'\bar{z}(c) \geqq b'z$ for any other feasible z, and hence $E \min_x C(b, x) \geqq Eb'z$ for any other feasible z. In particular, then, if z is the vector which maximizes $Eb'z$ subject to $z'A = c$, $z \geqq 0$, then $E \min_x C(b, x) \geqq Eb'z = \min_x C(Eb, x)$ by duality.

[2] E will denote the expectation with respect to the distribution of b throughout this paper.

[3] A simple proof of Jensen's inequality for continuous convex functions and arbitrary distributions is provided by noting that when the sets studied in Theorem 2 of [4] are unbounded, then the moment space R may be a proper subset of the convex hull D. But by a simple limiting process, Theorem 2 of [6] is the desired inequality.

We have thus proved the following theorem.

Theorem:

$$EC(b, \bar{x}(Eb)) \geqq \min_x EC(b, x) \geqq E \min_x C(b, x) \geqq \min_x C(Eb, x).$$

In particular, the inequality $\min_x EC(b, x) \geqq \min_x C(Eb, x)$ is the generalization of the observation in [5] that the expectation of the objective function for the optimal solution ($1,524,000) was greater than the minimum value of the objective function ($1,000,000) when the "right hand side" is replaced by its expectation.

It is well known that, when the probability measure on the set of b's is countably additive or the set of b's is with probability one finite in number, equality holds between $E \min_x C(b, x)$ and $\min_x C(Eb, x)$ if and only if $\min_x C(b, x)$ is a linear function of b (cf. [8], p. 265).

Trivially, if $\bar{x}(b) = \bar{x}$ for all b, or more generally if $C(b, x) = C(b, \bar{x}(b))$ except on a set of b of probability measure zero, then $\min_x EC(b, x) = E \min_x C(b, x)$.

A simple sufficient condition for equality between $\min_x EC(b, x)$ and $E \min_x C(b, x)$ is that $C(b, x)$ is a linear function of b, for then $\min_x EC(b, x) = \min_x C(Eb, x)$, and so by the above inequality $\min_x EC(b, x) = E \min_x C(b, x)$. From this observation the results of [9] and [10] become evident. For in these papers, the dynamic problem can be thought of as a succession of static "here-and-now" problems wherein $C(b, x)$ (in our notation) is of the form $\alpha'_1x + \alpha'_2b + x\alpha_3x' + b\alpha_4b' + x\alpha_5b' + b\alpha'_5x'$ where α_1 is an n_1-vector, α_2 is an m-vector, and α_3, α_4, and α_5 are $n_1 \times n_1$, $m \times m$, and $n_1 \times m$ matrices, respectively. Hence in minimizing $C(b, x)$ with respect to x, the quadratic term in b is not involved and so essentially $C(b, x)$ is a linear function of b. We see, then, that the x which minimizes $EC(b, x)$ is the x which minimizes $C(Eb, x)$, the result of [9] and [10]. In fact, we obtain the further result that if $C(b, x)$ can be expressed as $C_1(b, x) + C_2(b)$ where $C_2(b)$ involves only b (and not x) and $C_1(b, x)$ is a linear function of b, then use of Eb in place of the random b and solving the "certainty" linear programming problem also solves the "uncertainty" problem.[4]

For the "here-and-now" problem, the upper and lower bounds on $\min_x EC(b, x)$, namely $EC(b, \bar{x}(Eb))$ and $\min_x C(Eb, x)$, respectively, are easily computable. These bounds are also upper and lower bounds on $E \min_x C(b, x)$, the expected value of the objective function of the "wait-and-see" problem. We can, however, get an easily computable and sometimes sharper upper bound for $E \min_x C(b, x)$ by the following considerations.

[4] Reiter [7] studies the general problem of finding sufficient conditions for the solution of a stochastic problem to be that of a "surrogate" problem where the probability distribution is replaced by something simpler. His condition, in our case, is that $C(b, x) = \sum_{i=1}^n A_i(x)B_i(b)$, where $A_i(x) > 0$, $B_i(b) > 0$, in which case the x which minimizes $\sum_{i=1}^n A_i(x)EB_i(b)$ solves the stochastic problem. If we let $n = 3$, $B_1(b) = 1$, $B_2(b) = b$, $B_3(b) = C_2(b)$, and $A_3(x) = 1$, then $C(b, x) = C_1(b, x) + C_2(b) = A_1(x) + A_2(x)b + B_3(b)$, so that our result is a special case of Reiter's where the "surrogate" is Eb.

By using the theory of moment spaces of multivariate distributions, an upper bound on $Eg(X)$ has been obtained (cf. [6]), where $X' = (X_1, \cdots, X_r)$ is defined over an r-dimensional rectangle I_r (i.e., $-\infty < a_1 < a_2 < \infty$) and $g(X)$ is a continuous convex function of X. The upper bound has the following simple form when the X's are independent:

Let the 2^r vertices of I_r be written as $(a_{1\phi_1}, a_{2\phi_2}, \cdots, a_{r\phi_i})$ where ϕ_i $(i = 1, \cdots, r)$ takes on the values 1 and 2 and $a_{i1} < a_{i2}$ for all i. Then[5]

$$Eg(X) \leqq \sum_{\phi} \prod_{j=1}^{r} \frac{(-1)^{\phi_j}(a_{j\phi_j} - EX_j)}{(a_{j2} - a_{j1})} g(a_{1\bar{\phi}_1}, \cdots, a_{r\bar{\phi}_r}) = H^*(EX),$$

where $\bar{\phi}_i = 3 - \phi_i$.

We have shown above that $\min_x C(b, x)$ is a continuous convex function of b. The result just cited can then be rewritten as follows.

Theorem: If b is defined over a bounded m-dimensional rectangle $I_m(-\infty < \beta_1 < \beta_2 < \infty)$ and the b's are independent, then

$$E \min_x C(b, x) \leqq \sum_{\phi} \prod_{j=1}^{m} \frac{(-1)^{\phi_j}(\beta_{j\phi_j} - Eb_j)}{(\beta_{j2} - \beta_{j1})} \min_x C(\beta_{(\phi)}, x),$$
$$= H^*(Eb)$$

where $\beta_{(\phi)} = (\beta_{1\bar{\phi}_1}, \cdots, \beta_{m\bar{\phi}_m})$, ϕ_i takes on the values 1 and 2, $\bar{\phi}_i = 3 - \phi_i$ and ϕ is the set of 2^m m-dimensional vectors of 1's and 2's.

When the b's are not independent, a simple formula for the upper bound on $E \min_x C(b, x)$ based on the theory of moment spaces cannot be given. However, the upper bound can be computed as follows. Consider the class of sets of $m + 2$ points of the form $\beta_{(\phi)}$. Form the $S = \binom{2^m}{m+2}$ simplices in $m + 1$ dimensional space using each set of these $m + 2$ points as vertices. Let $H_s^+(Eb)$ be the point where the ray $(b_1 = Eb_1, \cdots, b_m = Eb_m, b_{m+1} = \theta \mid -\infty \leqq \theta \leqq \infty)$ pierces the upper boundary of the sth such simplex, $s = 1, \cdots, S$. Then

$$H^+(Eb) = \max_s H_s^+(Eb) \geqq E \min_x C(b, x).$$

A proof of this result is given in [6].

It has also been shown by Dantzig [3] (cf. also Theorem 2 of [1]) that $EC(b, x)$ is a convex function of x in a p-stage linear programming problem where only the right-hand side of the equation for the first stage is random. Let us now utilize this fact to obtain an easily computable lower bound for $\min_x EC(b, x)$ which is sometimes sharper than $E \min_x C(b, x)$. Since $EC(b, x)$ is convex in x, $\min_x EC(b, x) = EC(b, \bar{x})$ is greater than or equal to the value at $\bar{x}$ of a lower supporting hyperplane to the convex hull of $EC(b, x)$ which passes through any point $(x_1^0, \cdots, x_{n_1}^0, EC(b, x^0))$ on the curve $EC(b, x)$

[5] If some elements of X are non-random, without loss of generality assume they are the last p X's. Then the product runs to $r - p$ and $a_{i1} = a_{i2} = X_i$ for $i = r - p + 1, \cdots, r$.

(cf. [2]), where $x^0 = (x_1^0, \cdots, x_n^0)$. Then there are real numbers $\alpha_0^0, \alpha_1^0, \cdots, \alpha_{n_1}^0$ (depending on x^0) such that

$$\min_x EC(b, x) \geqq \alpha_0^0 + \sum_{j=1}^{n_1} \alpha_j^0 \bar{x}_j = H^0(\bar{x}),$$

say, where $EC(b, x^0) = \alpha_0^0 + \sum_{j=1}^{n_1} \alpha_j^0 x_j^0$. Take $x^0 = (\bar{x}(Eb))$, an easily determined point. Then if $EC(b, x)$ is differentiable at this point,

$$\alpha_j^0 = \left.\frac{\partial EC(b, x)}{\partial x_j}\right|_{x_j = \bar{x}_j(Eb)} \qquad j = 1, \cdots, n_1$$

and $\alpha_0^0 = EC(b, \bar{x}(Eb)) - \sum_{j=1}^{n_1} \alpha_j^0 \bar{x}_j(Eb)$, where the derivative is taken in the direction from x^0 to $\bar{x}$.

Hence, one can determine the coefficients $\alpha_0, \alpha_1 \cdots, \alpha_{n_1}$ without solving Dantzig's "here-and-now" problem and, if one has bounds on $\bar{x}$, can utilize this information to get a lower bound on $H^0(\bar{x})$. This then yields a lower bound on the value of the objective function of Dantzig's "here-and-now" problem which is independent of the solution of the problem.

Let us now look at Example 2 of [3] more closely with the above inequalities in mind. In this example,

$$x' = (x_{11}, x_{12}), \qquad y' = (x_{21}, x_{22}),$$

$$A = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}, \qquad B = \begin{pmatrix} 0 & 0 \\ 1 & -1 \end{pmatrix}, \qquad c' = (1, 0),$$

$$f' = (2, 0), \quad \text{and} \quad b' = (100, d_2).$$

In this case,

$$C(b, x) = \begin{cases} x_{11} & \text{if } x_{11} > d_2 \\ x_{11} + 2(d_2 - x_{11}) & \text{if } x_{11} \leqq d_2 \end{cases}.$$

If we consider the dual problem, we find that we must maximize $-100\, y_1 + d_2 y_2$, where the restraints on the y's can be illustrated geometrically by observing Figure 1.

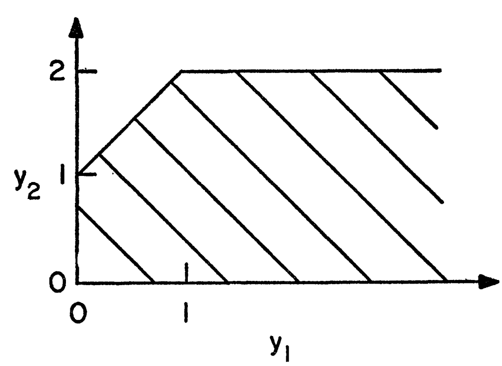

FIG. 1

The points (0, 0), (0, 1), or (1, 2) provide an optimal solution to the dual problem, depending on the value of d_2. Therefore in this case $\min_x C(b, x)$ is really a function of d_2, namely

$$\min_x C(b, x) = \begin{cases} 0 & \text{if } d_2 \leqq 0 \\ d_2 & \text{if } 0 < d_2 \leqq 100, \\ 2d_2 - 100 & \text{if } 100 < d_2 \end{cases}$$

which is clearly convex.

If d_2 has a uniform distribution between 70 and 80 (as in [3]), then $E_b \min_x C(b\ x) = Ed_2 = \min_x C(Eb, x) = 75$, since $\min_x C(b, x)$ is a linear function of b in this range. Also,

$$H^{+}(Eb) = H^{*}(Eb) = -\frac{(70 - Ed_2)}{(80 - 70)} 80 + \frac{(80 - Ed_2)}{(80 - 70)} 70 = Ed_2 = 75.$$

Since d_2 is assumed to be uniformly distributed between 70 and 80,

$$EC(b, x) = \begin{cases} -x_{11} + 150 & \text{if } x_{11} \leqq 70 \\ 77.5 + (75 - x_{11})^2/10 & \text{if } 70 < x_{11} \leqq 80. \\ x_{11} & \text{if } 80 < x_{11} \end{cases}$$

Note that $EC(b, x)$ is convex in x. This function attains its minimum, 77.5, when $x_{11} = 75$. Hence, $\min_x EC(b, x) = 77.5 = EC(b, \bar{x}(Eb))$, since $\bar{x}(Eb) = Ed_2 = 75$.

For example 2 of [3],

$$\frac{\partial EC(b, x)}{\partial x_{11}} = \begin{cases} -1 & \text{if } x_{11} \leqq 70 \\ \dfrac{x_{11} - 75}{5} & \text{if } 70 < x_{11} \leqq 80 \\ 1 & \text{if } 80 < x_{11} \end{cases}$$

and $\bar{x}_{11}(Eb) = 75$, so that $\alpha_1^0 = 0$ and the above inequality yields the inequality $\min_x EC(b, x) \geqq EC(b, \bar{x}(Eb))$. But since we already know that

$$EC(b, \bar{x}(Eb)) \geqq \min_x EC(b, x),$$

we see once again that equality holds between these values.

We see, then, that examination of the inequalities

$$H^0(\bar{x}) \leqq \min_x EC(b, x) \leqq EC(b, \bar{x}(Eb))$$

for the special "here-and-now" problem of Dantzig [3],

$$\min_x C(Eb, x) \leqq E \min_x C(b, x) \leqq H^{*}(Eb) \quad \text{or} \quad H^{+}(Eb)$$

for the "wait-and-see" problem, and the inequality which relates the "wait-and-see" problem to the two-stage "here-and-now" problem,

$$E \min_x C(b, x) \leqq \min_x EC(b, x),$$

yields information of use in evaluating whether or not the "expected value solution" is a good approximation to the solution of a linear programming problem under uncertainty.

References

1. Beale, E. M. L., "On Minimizing a Convex Function Subject to Linear Inequalities," *Journal of the Royal Statistical Society*, Series B, Vol. 17 (1955), pp. 173–184.
2. Bonnesen, T. and Fenchel, W., *Theorie der Konexen Körper*, Ergebnisse der Mathematik und ihrer Grenzgebiete, Vol. III, Part I, Berlin, J. Springer, 1934; reprinted Chelsea Publishing Co., New York, 1948.
3. Dantzig, G. B., "Linear Programming Under Uncertainty," *Management Science*, Vol. 1 (1955), pp. 197–206.
4. Dresher, M., Karlin, S. and Shapley, L. S., "Polynomial Games" Contributions to the Theory of Games, *Annals of Mathematics Study*, No. 24, 1950, pp. 161–180.
5. Ferguson, A. R., and Dantzig, G. B., "The Allocation of Aircraft to Routes—An Example of Linear Programming Under Uncertain Demand," *Management Science*, Vol. 3 (1956), pp. 45–73.
6. Madansky, A., "Bounds on the Expectation of a Convex Function of a Multivariate Random Variable," *Annals of Mathematical Statistics*, Vol. 30, (1959), pp. 743–746.
7. Reiter, S., "Surrogates for Uncertain Decision Problems: Minimal Information for Decision Making," *Econometrica*, Vol. 25 (1957), pp. 339–345.
8. Savage, L. J., *The Foundations of Statistics*, John Wiley and Sons, Inc., New York, 1954.
9. Simon, H. A., "Dynamic Programming Under Uncertainty with a Quadratic Criterion Function," *Econometrica*, Vol. 24 (1956), pp. 74–81.
10. Theil, H., "A Note on Certainty Equivalence in Dynamic Planning," *Econometrica*, Vol. 25 (1957), pp. 346–349.
11. Tintner, G., "Stochastic Linear Programming with Applications to Agricultural Economics," *Proceedings of the Second Symposium in Linear Programming*, National Bureau of Standards, Washington, D. C., Vol. I, 1955, pp. 197–228.
12. Vajda, S. "Inequalities in Stochastic Linear Programming," *Bulletin of the International Statistical Institute*, Vol. 36 (1958), pp. 357–363.