INEQUALITIES FOR STOCHASTIC NONLINEAR PROGRAMMING PROBLEMS

O. L. Mangasarian and J. B. Rosen*

Shell Development Company, Emeryville, California
(Received December, 1962)

Many actual situations can be represented in a realistic manner by the two-stage stochastic nonlinear programming problem $\text{Min}_x E\min_y[\varphi(x)+\psi(y)]$ subject to $g(x)+h(y)\geqq b$, where b is a random vector with a known distribution, and E denotes expectation taken with respect to the distribution of b. MADANSKY has obtained upper and lower bounds on the optimum solution to this two-stage problem for the completely linear case. In the present paper these results are extended, under appropriate convexity, concavity, and continuity conditions, to the two-stage nonlinear problem. In many cases of practical interest the calculation of these bounds will require only slightly more effort than two solutions of a deterministic problem of the same size, that is, a problem with a known constant value for the vector b. A small nonlinear numerical example illustrates the calculation of these bounds. For this example the bounds closely bracket the optimum solution to the two-stage problem.

IN THIS work† we obtain upper and lower bounds for the solution of the following two-stage stochastic nonlinear programming problem

$$\begin{aligned} &\text{Minimize}_x\ E\min_y[\varphi(x)+\psi(y)],\\ &\text{subject to } g(x)+h(y)\geqq b, \end{aligned} \tag{1}$$

where $\varphi(x)$ is a scalar function of the n_1-by-1 vector x, $\psi(y)$ is a scalar function of the n_2-by-1 vector y, $g(x)$ and $h(y)$ are each m-by-1 vectors whose components are scalar functions of their arguments, and b is a random m-by-1 vector with known distribution. E denotes the expectation taken with respect to the distribution of b. It should be noted that the constraints of problem (1) include any nonnegativity requirements on x and y.

Many actual situations can be represented in a realistic manner by the two-stage stochastic problem (1). One such example is that of a manu-

* Present address: Stanford University, Stanford, Calif.

† Vector notation will be used throughout this work. With obvious exception, lower case Roman letters will denote column vectors and Greek letters scalars. A prime will indicate the transpose. A subscript will denote a specific vector or scalar such as the ith vector b_i and the ith scalar γ_i. A second subscript will denote the component of the vector, thus b_{ij} will denote the jth component of the ith vector b_i.

facturing-inventory problem. The random vector b represents a demand for manufactured products that can only be specified in advance by its distribution. The vector x determines the amount of each product that will be made during some period before the actual demand is known. This first stage production of products corresponding to b is given by the vector $g(x)$ at a cost $\varphi(x)$. Once the actual demand b is known, we are free to choose the optimum value of the vector y that will compensate at minimum cost for shortages, that is, negative components of $g(x)-b$. This is the second stage, and results in a production or outside purchase vector $h(y)$ at a cost of $\psi(y)$. The problem is to select the first stage production vector x so as to minimize the expected value of the total cost $\varphi(x)+\psi(y)$.

Another example is a capital investment problem. Here, b may represent the required production rate during a typical period, again specified by its distribution only. The vector x determines the capital investment to achieve a production rate of $g(x)$. The value of $\varphi(x)$ is the appropriately discounted cost per period. When the actual required production rate is known, any shortages are met by an optimum choice of the vector y, which gives the additional production rate $h(y)$ at a cost of $\psi(y)$. The problem is to choose x so as to minimize the expected value of the discounted cost per period.

A linear 'two-stage' problem has previously been discussed by DANTZIG.[2] The results given in the next two sections are essentially generalizations of the inequalities obtained by MADANSKY[6] for the corresponding stochastic linear programming problem. The linear results are, of course, included in the results presented here.

For convenience, we define the function $\gamma(b, x)$ as follows:

$$\gamma(b,x)\equiv\min_y\{\varphi(x)+\psi(y)\,|\,g(x)+h(y)\geqq b\}.$$

The two-stage problem (1) can now be abbreviated as $\min_x E\gamma(b,x)$. When stated in this form it should be kept in mind that, in general, this represents a constrained minimization in the x-space. Thus, throughout this work we assume that there exists a convex set K such that for each x in K there is an associated y such that $g(x)+h(y)\geqq b$ for each b. We denote by 'feasible x' as those x in K.

Certain convexity and concavity properties are needed in what follows. We use the commonly accepted definition that $\varphi(x)$ is convex if, for $0\leqq\lambda\leqq 1$,

$$(1-\lambda)\varphi(x_1)+\lambda\varphi(x_2)\geqq\varphi[(1-\lambda)x_1+\lambda x_2]$$

for all x_1 and x_2 in the convex region of definition of $\varphi(x)$. The function $\varphi(x)$ is concave if the inequality sign is reversed.

The results presented here are of practical importance because they give an upper and a lower bound to the two-stage stochastic problem (1) in terms of the solutions for two *nonstochastic** problems of the *same size* as the original problem. In many cases the bounds obtained may be fairly close, as indicated by the numerical example in the last section. In order to obtain the lower bound (*see* Theorem 3) we solve the single deterministic problem in (n_1+n_2) variables and m inequalities,

$$\min_x \gamma(Eb,x) = \min_{x,y}\{\varphi(x)+\psi(y) \mid g(x)+h(y) \geqq Eb\}. \qquad (2)$$

We denote by $\bar{x}(Eb)$ the value of x which gives the minimum for (2). For any b with a finite distribution we obtain the upper bound given by Theorem 4 by solving a similar problem to (2) but with a different right-hand side vector, $b_{\max}$, as follows:

$$\min_x \gamma(b_{\max},x) = \min_{x,y}\{\varphi(x)+\psi(y) \mid g(x)+h(y) \geqq b_{\max}\}. \qquad (3)$$

The vector $b_{\max}$ is chosen so that $b_{\max} \geqq b$ for every b in the finite distribution.

An important class of problems is one where $\varphi(x)$ and $\psi(y)$ are convex functions and every component of $g(x)$ and $h(y)$ is a concave function of its variables. This class of problems is distinguished by the fact that for any fixed vector b the well-known necessary and sufficient Kuhn-Tucker conditions[4] for a global minimum apply. In particular, they are valid for the two problems (2) and (3). For such problems efficient computational methods have been developed. The gradient projection method[8] and convex partition programming[9] are two such methods that have been used to solve many problems of this class.

The original problem (1) can also be solved directly for the partially linear case when $\psi(y)$ and $h(y)$ are linear and b has a finite and discrete distribution, π_i, $i=1, \cdots, l$. This is done by introducing vectors y_i, $i=1, \cdots, l$, each of the same dimension as y. A formulation equivalent to (1) is given by the single deterministic problem with n_1+ln_2 variables and ml constraints,

$$\min_{x,y_i}\{\varphi(x)+\textstyle\sum_{i=1}^{i=l} \pi_i\psi(y_i) \mid g(x)+h(y_i) \geqq b_i\}. \qquad (i=1, \cdots, l) \quad (4)$$

Computational advantage is taken of the block diagonal structure of this formulation by noting that for any fixed x, the minimum with respect to the y_i is obtained by solving l *independent linear* subproblems. The convex partition programming algorithm[9] † is based on this structure and can be used to solve the problem (4). Several large problems have been solved successfully in this fashion on the IBM 7090 computer.

* The terms nonstochastic and deterministic will be used interchangeably. They refer to a problem for which the right-hand sides of the constraints are constants.

† The roles of x and y are interchanged in reference 9.

In the next section we derive some basic results that are needed to establish the desired inequalities, and in addition are of independent interest in connection with the parametric dependence of solutions to a nonlinear programming problem on its right-hand side (parametric nonlinear programming). The third section contains the main results and the final section gives a simple numerical example.

SOME BASIC RESULTS

WE BEGIN by establishing certain results, some, more general than needed for proving the inequalities of the next section, but which seem to be of sufficient interest to warrant their presentation. We define the following parametric nonlinear programming problem

$$\min_z \theta(z,a) \text{ subject to } f(z,a) \geqq 0,$$

where θ is a scalar function of the vectors z and a and f is a vector function of z and a. We establish now two basic lemmas.

LEMMA 1. *The scalar function* $\alpha(a) \equiv \min_z\{\theta(z,a) | f(z,a) \geqq 0\}$ *is a convex function of the vector a provided that* θ *is a convex function of the vector* $[z'a']$ *and each component of f is a concave function of* $[z'a']$.

Proof. Consider two arbitrary fixed values of a, a_1, and a_2, and let

$$\begin{aligned} \alpha(a_1) &= \min_z\{\theta(z,a_1) | f(z,a_1) \geqq 0\} \equiv \theta(z_1,a_1), \\ \alpha(a_2) &= \min_z\{\theta(z,a_2) | f(z,a_2) \geqq 0\} \equiv \theta(z_2,a_2). \end{aligned} \tag{5}$$

Note that (z_1,a_1) and (z_2,a_2) must satisfy

$$f(z_1,a_1) \geqq 0 \qquad \text{and} \qquad f(z_2,a_2) \geqq 0. \tag{6}$$

Now since every component of $f(z,a)$ is concave in $[z'a']$ it follows that for $0 \leqq \lambda \leqq 1$.

$$\begin{aligned} f[\lambda z_1 + (1-\lambda) z_2, \lambda a_1 + (1-\lambda) a_2 &\geqq \lambda f(z_1,a_1) + (1-\lambda) f(z_2,a_2) \\ &\geqq 0. \end{aligned} \qquad \text{(by 6)}$$

Whence the point $z = \lambda z_1 + (1-\lambda) z_2$ satisfies the constraint

$$f[z, \lambda a_1 + (1-\lambda) a_2] \geqq 0,$$

and hence

$$\begin{aligned} \min_z\{\theta[z,\lambda a_1 + (1-\lambda)a_2] | f[z,\lambda a_1 + (1-\lambda)a_2] \geqq 0\} \\ \leqq \theta[\lambda z_1 + (1-\lambda) z_2, \lambda a_1 + (1-\lambda) a_2]. \end{aligned} \tag{7}$$

Now

$$\begin{aligned} \alpha[\lambda a_1 + (1-\lambda)a_2] &= \min_z\{\theta[z,\lambda a_1 + (1-\lambda)a_2] | f[z,\lambda a_1 + (1-\lambda)a_2]\} \\ &\leqq \theta[\lambda z_1 + (1-\lambda)z_2, \lambda a_1 + (1-\lambda)a_2] && \text{(by 7)} \\ &\leqq \lambda\theta(z_1,a_1) + (1-\lambda)\theta(z_2,a_2) && [\text{by convexity of } \theta(z,a)] \\ &= \lambda\alpha(a_1) + (1-\lambda)\alpha(a_2). && \text{(by 5)} \end{aligned}$$

Hence

$$\alpha[\lambda a_1 + (1-\lambda)a_2] \leqq \lambda\alpha(a_1) + (1-\lambda)\alpha(a_2).$$

LEMMA 2. *The scalar function* $\alpha(a) \equiv \min_z\{\theta(z,a)(z,a) | f(z,a) \geqq 0\}$ *is a convex and continuous function of the vector a provided that* θ *is a convex and continuous function of the vector* $[z'a']$, *and each component of f is a concave and continuous function of* $[z'a']$.

Proof. The convexity of $\alpha(a)$ has been established in Lemma 1. It only remains to show its continuity. Let a_1 and a_2 be such that there exist z_1 and z_2 that satisfy $f(z_1, a_1) \geqq 0$ and $f(z_2, a_2) \geqq 0$. It follows from the concavity of f that for $0 \leqq \lambda \leqq 1$

$$f[\lambda z_1 + (1-\lambda)z_2, \lambda a_1 + (1-\lambda)a_2] \geqq \lambda f(z_1,a_1) + (1-\lambda)f(z_2,a_2)$$
$$\geqq 0.$$

Hence for $a = \lambda a_1 + (1-\lambda)a_2$ the value of $z = \lambda z_1 + (1-\lambda)z_2$ satisfies $f(z,a) \geqq 0$. It follows that the set of vectors a over which there exist feasible z is convex and hence the set of a's for which there exist $\alpha(a)$ is also convex. Since $\alpha(a)$ is a convex function over a convex set of a's it is continuous at every interior point of this convex set.

Now let a be on the boundary of this convex set and consider a sequence $\{a_i\}$ such that $\lim_{i=\infty} a_i = a$. Let z_i be the solution of $\min_z\{\theta(z,a_i) | \cdot f(z,a_i) \geqq 0\}$. Hence $f(z_i,a_i) \geqq 0$. Taking the limit superior* as i tends to infinity and invoking the continuity of f gives

$$f(\overline{\lim}_{i\to\infty} z_i, a) \geqq 0.$$

Hence

$$\min_z\{\theta(z,a) | f(z,a) \geqq 0\} \leqq \theta(\overline{\lim}_{i\to\infty} z_i, a),$$

and

$$\alpha(a) = \min_z\{\theta(z,a) | f(z,a) \geqq 0\}$$
$$\leqq \theta(\overline{\lim}_{i\to\infty} z_i, a)$$
$$= \overline{\lim}_{i\to\infty} \theta(z_i,a_i). \qquad \text{(by continuity of } \theta)$$

But

$$\alpha(a) \geqq \overline{\lim}_{i\to\infty} \theta(z_i,a_i);$$

hence

$$\alpha(a) = \overline{\lim}_{i\to\infty} \theta(z_i,a_i).$$

In a similar manner

$$\alpha(a) = \underline{\lim}_{i\to\infty} \theta(z_i,a_i).$$

Therefore $\alpha(a) = \lim_{i\to\infty} \theta(z_i,a_i)$.

THEOREM 1. *The scalar function* $\alpha(b) \equiv \min_x \gamma(b,x)$ *defined by*

$$\alpha(b) \equiv \min_x \gamma(b,x) = \min_x \min_y \{\varphi(x) + \psi(y) | g(x) + h(y) \geqq b\} \qquad (8)$$

* We denote the limit superior by $\overline{\lim}$ and the limit inferior by $\underline{\lim}$.

is a convex function of b, provided that $\varphi(x)$ *and* $\psi(y)$ *are convex functions and the components of* $g(x)$ *and* $h(y)$ *are concave functions of their respective arguments.*

Proof. Let $z=\begin{bmatrix} x \\ y \end{bmatrix}$, $\theta(z,a)=\varphi(x)+\psi(y), a=b, \alpha(a)=\alpha(b)$ and $f(z,a)=g(x)+h(y)-b$. We have only to show that $\theta(z, a)$ and $f(z, a)$ are respectively convex and concave in $[z'a']$ in order to invoke the Lemma 1 and prove the theorem. From the assumed concavity of $g(x)$ and $h(y)$ we have

$$g[\lambda x_1+(1-\lambda)x_2] \geqq \lambda g(x_1)+(1-\lambda)g(x_2),$$
$$h[\lambda y_1+(1-\lambda)y_2] \geqq \lambda h(y_1)+(1-\lambda)h(y_2),$$
$$-\lambda b_1-(1-\lambda)b_2=-\lambda b_1-(1-\lambda)b_2.$$

Addition of the last three relations and using the relations $f(z,a)=g(x)+h(y)-b$ and $a=b$ gives

$$f[\lambda z_1+(1-\lambda)z_2, \lambda b_1+(1-\lambda)b_2] \geqq \lambda f(z_1,b_1)+(1-\lambda)f(z_2,b_2),$$

establishing the concavity of $f(z, a)$ in the vector $[z'a']$. Similarly the convexity of $\theta(z, a)$ in the vector $[z'a']$ is established.

It is interesting to note that $\alpha(b)$ is also a nondecreasing function of b, that is $\alpha(b_1) \leqq \alpha(b_2)$ for $b_1 \leqq b_2$. This follows from the simple fact that if $b_1 \leqq b_2$ then $g(x)+h(y) \geqq b_2$ implies $g(x)+h(y) \geqq b_1$.

Theorem 1 is a generalization of Madansky's[6] Lemma 1 to the present nonlinear problem.

COROLLARY 1.* *The scalar function* $\alpha(b) \equiv \min_x \gamma(b,x)$ *defined by* (8) *is a continuous function of b provided that* $\varphi(x)$ *and* $\psi(y)$ *are convex and continuous functions and* $g(x)$ *and* $h(y)$ *are concave, and continuous functions of their respective arguments.*

Proof. This Corollary may be proven by using Theorem 1, or more simply by using Lemma 2. By letting $z=\begin{bmatrix} x \\ y \end{bmatrix}, \theta(z,a)=\varphi(x)+\psi(y)$, $a=b, \alpha(a)=\alpha(b)$ and $f(z,a)=g(x)+h(y)-b$; Lemma 2 may be invoked to establish the continuity of $\alpha(a)$ and hence that of $\alpha(b)$.

We proceed now to establish a theorem that is essentially a generalization of Theorem 2 of BEALE[1] and of a similar theorem by Dantzig.[2]

THEOREM 2. *The function* $\gamma(b,x)=\min_y\{\varphi(x)+\psi(y) \mid g(x)+h(y) \geqq b\}$ *is a convex function of the feasible x for any fixed b provided that* φ *and* ψ *are convex functions and the components of g and h are concave functions of their respective arguments.*

* We are indebted to ALBERT MADANSKY for pointing out the need for the continuity assumptions in this Corollary.

Proof. Define $a=x$, $z=y$, $\theta(z,a)=\varphi(x)+\psi(y)$, $f(z,a)=g(x)+h(y)-b$ and $\alpha(a)=\gamma(b,x)$. It can be easily shown (*see* the proof of Theorem 1 above) that the assumptions made in Theorem 2 on $\varphi(x)$, $\psi(y)$, $g(x)$, and $h(y)$ ensure the convexity of $\theta(z,a)$ and concavity of $f(z,a)$. Hence the Lemma may be invoked to establish the convexity of $\alpha(a)$ or the convexity of $\gamma(b, x)$ in x.

COROLLARY 2. *The function $E\gamma(b,x)$ is a convex function of the feasible x provided that the conditions of Theorem 2 are satisfied.*

Proof. Since $\gamma(b,x)$ is convex in x for a fixed b, integration over the distribution of b still gives a convex function of x, $E\gamma(b, x)$.

EXTENSION OF LINEAR RESULTS

IN THIS section we consider a number of additional problems, the results of which will serve as bounds on the original two-stage problem (1).

We mentioned earlier the nonstochastic problem of $\min_x\gamma(Eb,x)$. This is the problem in which the right side of the constraints is replaced by its expected value. Thus, for example, if we had a discrete distribution of the vectors b_i with probabilities π_i, $i=1, \cdots, l$, then this problem becomes

$$\min_x\gamma\left(\textstyle\sum_{i=1}^{i=l} \pi_i\, b_i,x\right).$$

We next introduce the 'wait-and-see' problem of Madansky,[6] $E\min_x(b,x)$. In this problem one waits for an observation of the random vector b and then solves the nonstochastic nonlinear programming problem based on this observed b. This is in contrast to Dantzig's[2] two-stage problem (1) where a decision x must be made at once without knowing b, and a subsequent decision y compensating for x is made after b is observed. We recall that in terms of the function $\gamma(b,x)$, the two-stage problem (1) is $\min_x E\gamma(b,x)$.

We are now in a position to derive one of the main results of this work, Theorem 3.

THEOREM 3. *Let $\bar{x}(Eb)$ be the solution of* $\min_x\gamma(Eb,x)$. *The inequalities*

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

hold provided that for the last inequality only, it is assumed that $\varphi(x)$ and $\psi(y)$ are convex and continuous functions and the components of $g(x)$ and $h(y)$ are concave and continuous functions of their respective arguments.

Proof. The first inequality obviously holds since the $\min_x E\gamma(b,x)$ is less than or equal to $E\gamma(b,x)$ evaluated as some x such as $\bar{x}(Eb)$.

For the second inequality let $\bar{x}$ be that solution which minimizes $E\gamma(b,x)$ and let $\bar{x}(b)$ be the solution which minimizes $\gamma(b,x)$. That is

$$\min_x(E\gamma(b,x)=E\gamma(b,\bar{x}), \quad \text{and} \quad E\min_x\gamma(b,x)=E\gamma[b,\bar{x}(b)].$$

Since for every b, $\gamma(b,\bar{x}) \geqq \gamma[b,\bar{x}(b)]$, then $E\gamma(b,\bar{x}) \geqq E\gamma[b,\bar{x}(b)]$ and hence

$$\min_x E\gamma(b,x) \geqq E\min_x \gamma(b,x),$$

which is the desired second inequality.

The third inequality is proved by using JENSEN's inequality,[3] (Theorem 2[5]), Theorem 1, and Corollary 1* of this paper. Jensen's inequality states that for a convex, continuous function β of a random vector z, $E\beta(z) \geqq \beta(Ez)$. Since by Theorem 1 and Corollary 1 $\min_x\gamma(b,x)$ is a convex and continuous function of b, it follows that

$$E\min_x\gamma(b,x) \geqq \min_x\gamma(Eb,x),$$

which is the desired third inequality.

For the case where b has a finite distribution, it is possible to get a somewhat rougher but easier-to-compute upper bound for the solution of the two-stage problem (1) by solving a single deterministic problem and obtaining a so-called 'fat solution.' This result follows from the following theorem.

THEOREM 4. *If the random vector b has a finite distribution, that is, $-\infty < b_{\min} \leqq b \leqq b_{\max} < \infty$, then the 'fat solution' $\min_x\gamma(b_{\max},x)$ is greater than or equal to the solution $\min_x E\gamma(b,x)$ of the two-stage problem.*

Proof. By definition we have

$$\min_x\gamma(b_{\max},x) = \min_x\min_y\{\varphi(x)+\psi(y) \mid g(x)+h(y) \geqq b_{\max}\}$$

and
$$\min_x E\gamma(b,x) = \min_x E\min_y\{\varphi(x)+\psi(y) \mid g(x)+h(y) \geqq b\}.$$

Now for any fixed x and b, every y that satisfies $g(x)+h(y) \geqq b_{\max}$ also satisfies $g(x)+h(y) \geqq b$. It follows then that $\gamma(b_{\max},x) \geqq \gamma(b,x)$ and thus $\gamma(b_{\max},x) \geqq E\gamma(b,x)$. Hence $\min_x\gamma(b_{\max},x) \geqq \min_x E\gamma(b,x)$.

Another upper bound for the 'wait-and-see' problem $E\min_x\gamma(b,x)$ can be derived by using the theory of moment spaces of multivariate distributions. The derivation of this bound is identical with Madansky, reference 6, p. 201, once Theorem 1 and Corollary 1 have been invoked to show that $\min_x(b,x)$ is a convex and continuous function of b.

THEOREM 5. *Provided that $\varphi(x)$ and $\psi(y)$ are convex and continuous functions, the components of $g(x)$ and $h(y)$ are concave and continuous functions of their arguments, b is defined over a bounded m-dimensional rectangle $b_1 \leqq b \leqq b_2 (-\infty < b_1 < b_2 < \infty)$ and the b's are independent, then*

$$E\min_x\gamma(b,x) \leqq \sum_{i=1}^{i=2^m}\{\textstyle\prod_{j=1}^{j=m} [(-1)^{f_{ij}}(b_{f_{ij},j} - (Eb)_j]/(b_{2j} - b_{1j})\} \cdot$$
$$\min_x\gamma(b_{3-f_{i1},1}, \cdots, b_{3-f_{im},m}, x),$$

where f_i is an m-dimensional vector $[f_{i1}, f_{i2}, \cdots, f_{im}]'$ whose components are all

* This inequality can also be proved directly, but under somewhat more restrictive conditions, by using the duality theorems of nonlinear programming.[10,7]

1's and/or 2's and the set of 2^m vectors f_i, $i=1, \cdots, 2^m$ is the set of all possible arrangements of 1's and/or 2's taken m at a time, and $(Eb)_j$ is the expected value of the jth component of b.*

Note that in order to calculate the above bound one needs to solve in general 2^m nonlinear programming problems, which may be a fairly lengthy procedure.

We now state and prove our last bound, a lower bound for the two-stage problem.

THEOREM 6. *Provided that φ and ψ are convex functions and the components of g and h are concave functions and $E\gamma(b,x)$ is differentiable at $x=\bar{x}(Eb)$ then*

$$\min_x E\gamma(b,x) \geqq E\gamma[b,\bar{x}(Eb)]+[\bar{x}-\bar{x}(Eb)]'\nabla E\gamma[b,\bar{x}(Eb)],$$

where $\bar{x}(Eb)$ is the solution that minimizes $\gamma(Eb,x)$, $\bar{x}$ is the solution that minimizes $E\gamma(b,x)$, and ∇ is a column vector of partial differential operators: $(\partial/\partial x_1, \cdots, \partial/\partial x_{n_2})$.

Proof. By Corollary 2, $E\gamma(b,x)$ is a convex function of x. Now if a convex function $\theta(x)$ is differentiable at some point $x=x_1$ it follows from the definitions of convexity and differentiability that (reference 4, p. 485)

$$\theta(x_2)-\theta(x_1) \geqq (x_2-x_1)'\nabla\theta(x_1).$$

Now if we let $\theta(x)=E\gamma(b,x)$ $x_2=\bar{x}$ and $x_1=\bar{x}(Eb)$ we have

$$E\gamma(b, \bar{x}) \geqq E\gamma[b,\bar{x}(Eb)]+[\bar{x}-\bar{x}(Eb)]'\nabla E\gamma[b,\bar{x}(Eb)],$$

which is precisely the assertion of Theorem 6 if we recall that $E\gamma(b,\bar{x})=\min_x E\gamma(b,x)$.

Note that in order to use the bound given by the above theorem one has to have bounds on $\bar{x}$ that appear on the right-hand side. However, if either $\bar{x}=\bar{x}(Eb)$ or if $\nabla E\gamma[b,\bar{x}(Eb)]=0$, then it follows from this Theorem and the first inequality of Theorem 3 that $E\gamma[b,\bar{x}(Eb)]=\min_x E\gamma(b,x)$.

NUMERICAL EXAMPLE

WE TAKE Example 2 of Dantzig,[2] change the linear cost function $x+2y$ to the nonlinear cost x^2+2y^2 and eliminate the slack variables. We thus consider the following problem in the scalar quantities x, y, and b.

$$\begin{aligned} &\text{Minimize}_x \quad E\min_y(x^2+2y^2),\\ &\text{subject to}\\ &\qquad -x\geqq -100,\\ &\qquad x+y\geqq b,\\ &\qquad x\geqq 0,\\ &\qquad y\geqq 0,\end{aligned}$$

* For example for $m=2$, $f_1=(1\ 1)'$, $f_2=(1\ 2)'$, $f_3=(2\ 1)'$, $f_4=(2\ 2)'$.

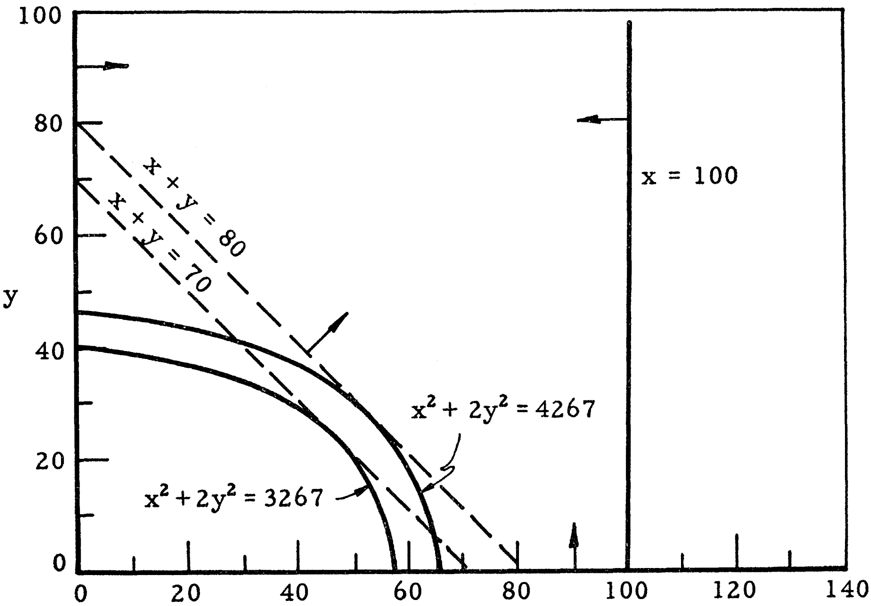

Figure 1

where b is an unknown demand uniformly distributed between 70 and 80. Figure 1 is a graphical representation of the problem. The dotted lines $x+y=70$ and $x+y=80$ bound the feasible region when $b=70$ and $b=80$ respectively. It is easy to verify that this example satisfies all the restrictions imposed by the various theorems and corollaries of this work. This example illustrates that the inequalities can be used to obtain an accurate estimate of $\min_x E\gamma(b,x)$ in a nontrivial case.

We begin first by determining the function $\gamma(b,x)$, which is the $\min_y(x^2+2y^2)$ subject to all the constraints. After some simple considerations with the aid of Fig. 1 we conclude that

$$\gamma(b,x)=\begin{cases} x^2 & \text{for} \quad 100\geqq x\geqq b;(y=0), \\ x^2+2(b-x)^2 & \text{for} \quad b\geqq x\geqq 0;(y=b-x). \end{cases} \tag{9}$$

To determine $\min_x\gamma(Eb,x)$, we replace b in (9) by its expected value of 75 and find the minimum of $\gamma(75,x)$, which turns out to be

$$\min_x\gamma(Eb,x)=3750 \quad \text{at} \quad x=\bar{x}(Eb)=50. \tag{10}$$

By considering (9) and Fig. 1 it is possible to determine

$$\min_x \gamma(b,x)=\begin{cases} 0 & \text{for } 0\geqq b;(x=0,y=0), \\ \tfrac{2}{3}b^2 & \text{for } 150\geqq b\geqq 0; \\ & \quad (x=\tfrac{2}{3}b,y=\tfrac{1}{3}b), \\ 10{,}000+2(b-100)^2 & \text{for } b\geqq 150; \\ & \quad (x=100,y=b-100). \end{cases} \tag{11}$$

Hence

$$E\min_x \gamma(b,x)=3756. \tag{12}$$

To determine $E\gamma(b,x)$ we have to consider three cases: $70\geqq x\geqq 0$, $80\geqq x\geqq 70$ and $100\geqq x\geqq 80$. The first and third cases are straightforward and $E\gamma(b,x)$ is obtained from (9) by $E[x^2+2(b-x)^2]$ and $E[x^2]$ respectively. For the second case, $80\geqq x\geqq 70$, we have from (9)

$$E\gamma(b,x)=\int_{b=70}^{b=x} x^2\frac{db}{10}+\int_{b=x}^{b=80} x^2+2(b-x)\frac{db}{10}.$$

Carrying the above calculations results in

$$E\gamma(b,x)=\begin{cases} 3x^2-300x+\tfrac{1}{3}(33800) & \text{for } 70\geqq x\geqq 0, \\ \tfrac{1}{10}[-\tfrac{2}{3}x^3+170x^2-12800x \\ \quad +\tfrac{1}{3}(1024000)] & \text{for } 80\geqq x\geqq 70, \\ x^2 & \text{for } 100\geqq x\geqq 80. \end{cases} \tag{13}$$

From this it is easy to conclude that

$$\min_x E\gamma(b,x)=3767 \quad \text{at} \quad x=\bar{x}=50. \tag{14}$$

From (10) and (13) we determine that

$$E\gamma[b,\bar{x}(Eb)]=3767. \tag{15}$$

Hence the inequalities of Theorem 3 are satisfied by the results given by (15), (14), (12), and (10), that is

$$3767=3767>3756>3750,$$

which shows that the easily calculated first and last quantities give a good estimate of the middle two, in a nontrivial example.

We now compute $\min_x\gamma(b_{\max},x)=\min_x\gamma(80,x)$. From (9) and Fig. 1 we see that $\min_x\gamma(b_{\max},x)=\gamma(80,53.3)=4267$. As asserted by Theorem 4 this is larger than $E\min_x\gamma(b,x)=3756$.

We also note that by considering various first and second partial derivatives of $\gamma(b,x)$, $\min_x\gamma(b,x)$ and $E\gamma(b,x)$ as given by (9), (11), and (13) we conclude that $\min_x\gamma(b,x)$ is a convex and continuous function of b, $\gamma(b,x)$ is convex in x for a fixed b, and that $E\gamma(b,x)$ is a convex function of x. These indeed are the assertions of Theorem 1, Corollary 1, Theorem 2, and Corollary 2.

We next compute the expression given in Theorem 5 with the help of (11) as follows

$$[2(75-70)(80)\tfrac{2}{3}(80-70)]-[2(75-80)(70)\tfrac{2}{3}(80-70)]=3767,$$

which is indeed greater than $E\min_x \gamma(b,x)$.

We finally turn to Theorem 6. From (10) and (14) we have $\bar{x}(Eb)=\bar{x}$. Hence by Theorem 6, $\min_x E\gamma(b,x) \geqq E\gamma(b,\bar{x}(Eb))$. But by Theorem 3 the reverse inequality must also hold, hence $\min_x E\gamma(b,x)=E\gamma(b,\bar{x})$, which is indeed the case as observed from (14) and (15).

REFERENCES

1. E. M. L. Beale, "On Minimizing a Convex Function Subject to Linear Inequalities," *J. Roy. Stat. Soc. B* **17**, 173–184 (1955).
2. G. B. Dantzig, "Linear Programming Under Uncertainty," *Management Sci.* **1**, 197–206 (1955).
3. J. L. W. V. Jensen, "Sur les fonctions convexes et les inégalités entre les valeurs moyennes," *Acta Math.* **30**, 175–193 (1906).
4. H. W. Kuhn and A. W. Tucker, "Nonlinear Programming," *Proc. Second Berkeley Symposium in Math. Stat. and Probability*, pp. 481–492, Aug. 1950, University of California Press, 1951.
5. A. Madansky, "Bounds on the Expectation of a Convex Function of a Multivariate Random Variable," *Ann. Math. Stat.* **30**, 743–746 (1959).
6. ———, "Inequalities for Stochastic Linear Programming Problems," *Management Sci.* **6**, 197–204 (1960).
7. O. L. Mangasarian, "Duality in Nonlinear Programming," Paper P-1054, Shell Development Company, Emeryville, California, 1961; *Quart. Applied Math.* **20**, 300–302 (1962).
8. J. B. Rosen, "The Gradient Projection Method for Nonlinear Programming, Part I. Linear Constraints," *J. Soc. Ind. Appl. Math.* **8**, 181–217 (1960); "Part II. Nonlinear Constraints," **9**, 514–532 (1961), (Program available as SHARE distribution No. 1399).
9. ———, "Convex Partition Programming," *Symposium on Math. Programming*, University of Chicago, June 1962; *Advances in Mathematical Programming*, L. M. Graves and P. Wolfe (eds.), McGraw-Hill, 1963.
10. Philip Wolfe, "A Duality Theorem for Nonlinear Programming," *Quart. Applied Math.* **19**, 239–244 (1961).