

BOUNDING SEPARABLE RECOURSE FUNCTIONS WITH LIMITED DISTRIBUTION INFORMATION *

John R. BIRGE

Industrial and Operations Engineering, University of Michigan, Ann Arbor, MI 48109, USA

José H. DULÁ

Operations Research Program, Southern Methodist University, Dallas, TX 75275, USA

The recourse function in a stochastic program with recourse can be approximated by separable functions of the original random variables or linear transformations of them. The resulting bound then involves summing simple integrals. These integrals may themselves be difficult to compute or may require more information about the random variables than is available. In this paper, we show that a special class of functions has an easily computable bound that achieves the best upper bound when only first and second moment constraints are available.

Keywords: Integration, stochastic programming, moment problem, duality, approximation.

1. Introduction

The recourse problem in stochastic programming is to find the expected value of minimizing the cost of meeting some set of constraints that may depend on a random variable, ω . In this analysis, we suppose that the resulting function is convex. In general, we seek to bound the integral of this or any convex function which is too expensive for numerical integration or, for which, only limited distributional information is known. Thus, many problems in applied mathematics, such as calculating the average load on a structure, the present value of a stock option, or the expected performance of a computer system, fit this framework.

The general problem is to find:

$$E_f(x) = E\{f(x(\omega))\} = \int f(x(\omega))P(d\omega), \quad (1.1)$$

* This research has been partially supported by the National Science Foundation under Grants ECS-8304065 and ECS-8815101, by the Office of Naval Research Grant N00014-86-K-0628 and by the National Research Council under a Research Associateship at the Naval Postgraduate School, Monterey, California.

where x is a random vector mapping the probability space, (Ω, \mathcal{A}, P) , onto $(\mathbb{R}^N, \mathcal{B}^N, F)$, F is the distribution function of x , and $x \in X \subset \mathbb{R}^N$. The *expectation functional*, $E_f(x)$ can also be written as a Lebesgue–Stieltjes integral with respect to F :

$$E_f(x) = \int_{\mathbb{R}^N} f(x) \, dF(x). \quad (1.2)$$

Difficulties arise in evaluating $E_f(x)$ when either the function f is difficult to evaluate or the distribution function F is not known exactly. In stochastic programming, the function f is the optimal value of a recourse action which depends on $x(\omega)$. This function is convex for the following recourse problem in which the random vector appears as a linear term in the constraints. In this case,

$$f(x(\omega)) = \inf_{y(\omega) \in \mathbb{R}^m} \{q(y(\omega)) \mid g(y(\omega)) \leq x(\omega) \text{ a.s.}\}, \quad (1.3)$$

where $q: \mathbb{R}^m \rightarrow \mathbb{R}$ and $g: \mathbb{R}^m \rightarrow \mathbb{R}^N$ are convex. In general, each evaluation of f requires the solution of a mathematical program. Although many approximation formulas for integrals (1.2) have been given (see Davis and Rabinowitz [8]), the expense in these computations and the possible high values of N make them inefficient.

In this paper, we concentrate on approximations that first bound f with a separable function and then compute integrals in each variable separately. We suppose that only first, second moment and some range information is available and show that a tight bound on the expectation of a large class of convex functions with known first and second moments can be obtained in a single linesearch. This result extends results for linear subproblems (1.3) and contrasts with previous methods that either require generalized programming [3,6,14] or obtain a looser bound through linear approximation [5].

Section 2 provides background on previous approaches to bounding expectations. The generalized moment problem interpretation is given in section 3. Section 4 presents the separable approximation used for the bound. Section 5 presents basic results for bounds on each separable component and some examples of the general class of functions allowed in this analysis.

2. Background and previous approximations

Many of the approximation schemes for stochastic programs with recourse are described in Birge and Wets [3] and Kall [24]. For general functions f , the basic procedures to approximate $E_f(x)$ use some form of a discrete approximation for the distribution of x . Numerical integration procedures are often based on the *midpoint* and the *trapezoidal* approximations. On an interval, $[a, b]$, the approximations are improved by dividing $[a, b]$ into subintervals, appropriately

weighting the subintervals and applying the midpoint and trapezoidal approximations on each subinterval.

A more sophisticated procedure is *Gaussian quadrature* to find an integral formula that fits all polynomials up to some degree. As noted by Miller and Rice [33], this can be used to find a discretization with K values that matches the first $K + 1$ moments of the distribution of x . To match the first three moments of the uniform distribution on $[a, b]$, for example, Gaussian quadrature selects two points, $(a + b)/2 \pm (\sqrt{3}/6)(b - a)$, with equal probability, $1/2$.

A difficulty with using the Gaussian quadrature formulas is that they do not generally provide bounds on the expectation. Restrictions on higher-order derivatives and Peano’s theorem [34] provide bounds but they require, at least, differentiability of f and a density function that may not be available. Generalizations of the midpoint and trapezoidal approximations do, however, obtain bounds on the expectation of a convex function. For example, Jensen’s inequality [23] can be interpreted as a generalization of the midpoint approximation that provides a lower bound on the expected value of convex f through:

$$\int_{\mathbb{R}^N} f(x) \, dF(x) \geq f\left(\int_{\mathbb{R}^N} x \, dF(x)\right), \tag{2.3}$$

where $\bar{x} = \int_{\mathbb{R}^N} x \, dF(x)$ is assumed finite.

Madansky, following Edmundson [13,31], provided a generalization of the trapezoidal approximation, called the *Edmundson–Madansky inequality*, that gives an upper bound on the expectation of a convex function. For $N = 1$, the basic inequality is:

$$\int_{\mathbb{R}^N} f(x) \, dF(x) \leq \frac{((b - \bar{x})f(a) + (\bar{x} - a)f(b))}{(b - a)}, \tag{2.4}$$

where $X = [a, b]$. The Edmundson–Madansky inequality (2.4) can also be extended to multiple dimensions and infinite intervals (see, for example, [1,15,18]).

Refinements of the Jensen and Edmundson–Madansky inequalities are possible by subdividing the interval (or, more generally, the region) into smaller pieces on which the bounds can be reapplied as in the traditional midpoint and trapezoidal approximations (see [3,17,21,26]). These refinements require additional functional evaluations and conditional expectations on the subregions. As has been observed, the Jensen lower bound is generally reasonably accurate relative to the Edmundson–Madansky upper bound (e.g., [19]), which requires a number of function evaluations that increases exponentially in the number of random variables. The primary concern is then in obtaining more accurate upper bounds without additional computational effort.

A bound for linear recourse problems that requires linear work in the number of the random variables was introduced in [3] and extended in [4] and [2]. A similar bound also appears for network recourse problems in Wallace [40]. Related extensions to functions built on simplicial decomposition of the function

also appear in [10] and [16]. An iterative approach for piecewise linear recourse appears in [6].

This paper builds on the idea of introducing separability into the recourse function. It extends this idea of separability to the nonlinear recourse problem in (1.3) and assumes only that first and second moment information is known about each of the random components. This is especially important when the random components are transformations of some set of original random variables.

The results here show that bounds on each component are easily computable in a single linesearch for a broad class of convex functions, if second moment information is available. We note that these results can also be seen as extensions of Kall's result [25] in this volume for bounding the optimal value of the linear recourse problem with only first and second moment information. In both analyses, a generalized moment problem formulation leads to the bound. This problem is described in the next section.

3. Generalized moment problem

To obtain bounds that hold for all distributions with certain properties, we can find

$Q \in \mathcal{P}$ a set of probability measures on (X, \mathcal{B}^N) subject to

$$\int_X v_i(x) Q(dx) \leq \alpha_i, \quad i = 1, \dots, s, \tag{3.1}$$

$$\int_X v_i(x) Q(dx) = \alpha_i, \quad i = s + 1, \dots, M,$$

to maximize $\int_X f(x) Q(dx)$,

where M is finite and the v_i are bounded, continuous functions. A solution of (3.1) obtains an upper bound on the expectation of f with respect to any probability measure satisfying the conditions above. Problem 3.1 is a *generalized moment problem* [29]. When the v_i are powers of x , the constraints restrict the moments of x with respect to Q . In this context, (3.1) determines an upper bound when only limited moment information on a distribution is available.

Problem 3.1 can also be interpreted as an *abstract linear program* since the objective and constraints are linear functions of the probability measure. The solution is then an extreme point (see [37] for a discussion of properties) in the infinite dimensional space of probability measures. The following theorem, proven in [27, theorem 2.1], gives the explicit solution properties.

THEOREM 3.1

Suppose X is compact. Then the set of feasible measures in (3.1), \mathcal{Q} , is convex and compact (with respect to the weak* topology), and \mathcal{Q} is the closure of the

convex hull of the extreme points of \mathcal{Q} . If f is continuous relative to X , then an optimum (maximum or minimum) of $\int_X f(x)Q(dx)$ is attained at an extreme point of \mathcal{Q} . The extremal measures of \mathcal{Q} are those measures that have finite support, $\{x_1, \dots, x_L\}$, with $L \leq M + 1$, such that the vectors,

$$\left(\begin{matrix} v_1(x_1) \\ v_2(x_1) \\ \vdots \\ v_M(x_1) \\ 1 \end{matrix} \right), \dots, \left(\begin{matrix} v_1(x_L) \\ v_2(x_L) \\ \vdots \\ v_M(x_L) \\ 1 \end{matrix} \right), \tag{3.2}$$

are linearly independent.

Kemperman [28] showed that the supremum is attained under more general continuity assumptions and provides conditions for \mathcal{Q} to be nonempty. Dupačová (formerly Žáčková) [11,12,41] pioneered the use of the moment problem as a bounding procedure for stochastic programs in her work on a minimax approach to stochastic programming. She showed that (3.1) attains the Edmundson–Madansky bound (and the Jensen bound if the objective is minimized) when the only constraint in (3.1) is $v_1 = x$, i.e., the constraints fix the first moment of the probability measure. She also provided some properties of the solution with an additional second moment constraint ($v_2(x) = x^2$) for a specific objective function f . To solve (3.1) generally, we consider a *generalized linear programming* procedure (see [7, chapter 24]).

GENERALIZED LINEAR PROGRAMMING PROCEDURE FOR THE GENERALIZED MOMENT PROBLEM (GLP)

Step 0. Initialization. Identify a set of $L \leq M + 1$ linearly independent vectors as in (3.2) that satisfy the constraints in (3.1). (Note that a *phase-one objective* [7] may be used if such a starting solution is not immediately available. For $N = 1$, the Gaussian quadrature points may be used as mentioned above.) Let $\nu = L$, $k = 1$, go to 1.

Step 1. Master problem solution. Find $p_1 \geq 0, \dots, p_\nu \geq 0$ such that

$$\begin{aligned} \sum_{l=1}^{\nu} p_l &= 1, \\ \sum_{l=1}^{\nu} v_i(x_l) p_l &\leq \alpha_i, \quad i = 1, \dots, s, \\ \sum_{l=1}^{\nu} v_i(x_l) p_l &= \alpha_i, \quad i = s + 1, \dots, M, \quad \text{and} \\ z &= \sum_{l=1}^{\nu} f(x_l) p_l \text{ is maximized.} \end{aligned} \tag{3.3}$$

Let $\{p_1^k, \dots, p_\nu^k\}$ attain the optimum in (3.3), and let $\{\theta^k, \pi_1^k, \dots, \pi_M^k\}$ be the associated *dual multipliers* such that

$$\begin{aligned} \theta^k + \sum_{i=1}^M \pi_i^k v_i(x_l) &= f(x_l), \quad \text{if } p_l^k > 0, \quad l = 1, \dots, \nu, \\ \theta^k + \sum_{i=1}^M \pi_i^k v_i(x_l) &\geq f(x_l), \quad \text{if } p_l^k = 0, \quad l = 1, \dots, \nu, \\ \pi_i^k &\geq 0, \quad i = 1, \dots, s. \end{aligned} \tag{3.4}$$

Step 2. Subproblem solution. Find $x^{\nu+1}$ that maximizes

$$\rho(x, \theta^k, \pi^k) = f(x) - \theta^k - \sum_{i=1}^M \pi_i^k v_i(x). \tag{3.5}$$

If $\rho(x^{\nu+1}, \theta^k, \pi^k) > 0$, let $\nu = \nu + 1$, $k = k + 1$ and go to 1. Otherwise, stop, $\{p_1^k, \dots, p_\nu^k\}$ are optimal probabilities associated with $\{x_1, \dots, x_\nu\}$ in a solution to (3.1).

The proof of the convergence of GLP is given in [7, chapter 24]. This result is used in [14] to solve a class of problems (3.1). The difficulty in GLP is in the solution of the subproblem (3.5), which generally involves a nonconvex function. Birge and Wets [3] describe how to solve (3.5) with constrained first and second moments, if convexity properties of ρ can be identified. Cipra [6] describes other methods for this problem based on discretizations and random selections of candidate points, x_i . Dulá [10] gives results when f is sublinear and has simplicial level sets. Kall [25] gives the results for sublinear, polyhedral functions with known generators.

In this paper, we first develop a separable bounding approximation for f as defined in (1.3). We then give conditions so that (3.1) can be solved for this separable function without requiring the repeated nonconvex optimization in (3.5). We show that bounds can be obtained for a general recourse problem by bounding several functions in one dimension that each require only $L = 2$ points of support that can be identified in one line search. This result gives bounds on the expectation of f with only $O(N)$ function evaluations.

4. Separable bounds in the general recourse problem

The use of the generalized programming formulation is limited in multiple dimensions because of the difficulty in solving subproblem (3.5). These computational disadvantages for large values of N suggest that a looser but more computationally efficient upper bound on the value of (3.1) may be more useful than solving (3.1) exactly for large N .

If a separable function, $v(x) = \sum_{i=1}^N v_i(x(i))$, is available, it offers an obvious advantage by only requiring single integrals. In this case, we would like to find $v(x) = \sum_{i=1}^N v_i(x(i)) \geq f(x)$ where each $v_i(x(i))$ is a convex function. Methods for constructing these functions to bound the optimal value of a linear program with random right-hand side are discussed in [2] and [4]. We give below the results for the general problem in (1.3).

LEMMA 4.1

If f is defined as in (1.3), then f is a convex function of x .

Proof

Let y_1 solve the optimization problem in (1.3) for x_1 and let y_2 solve the corresponding problem for x_2 . Consider $x = \lambda x_1 + (1 - \lambda)x_2$. In this case, $g(\lambda y_1 + (1 - \lambda)y_2) \leq \lambda g(y_1) + (1 - \lambda)g(y_2) \leq \lambda x_1 + (1 - \lambda)x_2$. So, $f(\lambda x_1 + (1 - \lambda)x_2) \leq q(\lambda y_1 + (1 - \lambda)y_2) \leq \lambda f(x_1) + (1 - \lambda)f(x_2)$, giving the result. \square

Let

$$v_i(x(i)) \equiv \frac{1}{N} f(Nx(i)e_i), \tag{4.1}$$

which is the optimal value of a parametric mathematical program. The following theorem shows that these values supply the separable bound required. Related bounds are possible by defining v_i with other right scalar multiples, $f\lambda_i(x(i)e_i)$ (see [36] for general properties), where $\sum_{i=1}^N \lambda_i = 1$. The proof below is easily extended to these cases and to translations of the constraints and explicit variable bounds as in [2] and [40].

THEOREM 4.1

The function $v(x) = \sum_{i=1}^N v_i(x(i)) \geq f(x)$, where f is defined as in (1.3).

Proof

In this case, let $y_i(x(i))$ solve (1.3), where $x(\omega) = Nx(i)e_i$. Then,

$$g\left(\frac{\sum_{i=1}^N y_i(x(i))}{N}\right) \leq \sum_{i=1}^N \left(\frac{1}{N}\right) [g(y_i(x(i)))] \leq \sum_{i=1}^N \left(\frac{1}{N}\right) Nx(i)e_i = x.$$

Next, let y^* solve (1.3) for x in the right-hand side of the constraints. By feasibility of $\sum_{i=1}^N y_i(x(i))/N$,

$$\begin{aligned} f(x) &= q(y^*) \leq q\left(\sum_{i=1}^N \frac{y_i(x(i))}{N}\right) \leq \sum_{i=1}^N \left(\frac{1}{N}\right) q(y_i(x(i))) = \sum_{i=1}^N v_i(x(i)) \\ &= v(x). \quad \square \end{aligned}$$

This results demonstrates that a parametric optimization of (1.3) in $i = 1, \dots, N$ yields an upper bound on $f(x)$ for any x . The bound may be tight as in some examples for stochastic linear programs as given in [4]. In these problems, $f(x) = \min_{y \in \mathbb{R}^n} \{q^T y \mid Ay = x\}$. Note that f is then a sublinear function and that $v_i(x(i)) = f(x(i)e_i)$. The functions v_i are found by solving for $q_i^\pm = \min_{y \in \mathbb{R}^n} \{q^T y \mid Ay = \pm e_i\}$ and letting

$$v_i(x(i)) = \begin{cases} q_i^+ x(i) & \text{if } x(i) \geq 0, \\ -q_i^- x(i) & \text{if } x(i) < 0. \end{cases} \tag{4.2}$$

Generalizations of the stochastic linear program bound as in [4] can also be given for the general bound in theorem 4.1. For example, we may apply a linear transformation T to x to obtain $u = Tx$. The constraints become $g(y) \leq T^{-1}(u)$. A new bound may be found by letting $\mu(u(i)) = (1/N) \inf\{q^T y \mid T(g(y)) \leq N(T_i^{-1}u(i))\}$. In linear problems, the convex hull of several of these bounds obtains good overall approximations (see [4]). We note that adjusting $1/N$ to other multipliers λ_i may also lead to better bounds.

To use any bound of the general type in theorem 3.1 to bound $\int_{\mathbb{R}^N} f(x) dF(x)$ requires a bound on $\int_{\mathbb{R}} v_i(x(i)) dF_i(x_i)$ or $\int_{\mathbb{R}} \mu_i(u(i)) dF_{u_i}(u(i))$, where F_i is the marginal distribution on x_i and F_{u_i} is the marginal distribution on $u(i)$. Since function evaluations may be expensive (solving (1.3)) and distribution information may be limited (especially in the case of F_{u_i}), we use the generalized moment problem to obtain bounds on each integral in \mathbb{R} . Generalized linear programming may solve this problem but it can be inefficient. In the next section, we show that a large class of functions require only two points of support in the bounding distribution. A single linesearch can determine these points and give a bound on f over all distributions with bounded first and second moments for the marginals.

5. Two-point support functions

To ease the notation in this section, we develop bounds on $\int v_i(x(i)) dF_i(x(i))$ by referring to f as a function on \mathbb{R} ($N = 1$). We then consider the moment problem (3.1) with $s = 0$, and $M = 2$ and where the constraints correspond to known first and second moments. In other words, we wish to find:

$$\begin{aligned} U &= \sup_{Q \in \mathcal{P}} \int_X f(x) Q(dx) \\ \int_X x Q(dx) &= \bar{x} \\ \int_X x^2 Q(dx) &= x^{(2)}, \end{aligned} \tag{5.1}$$

where \mathcal{P} is the set of probability measures on (X, \mathcal{B}^1) , the first moment of the

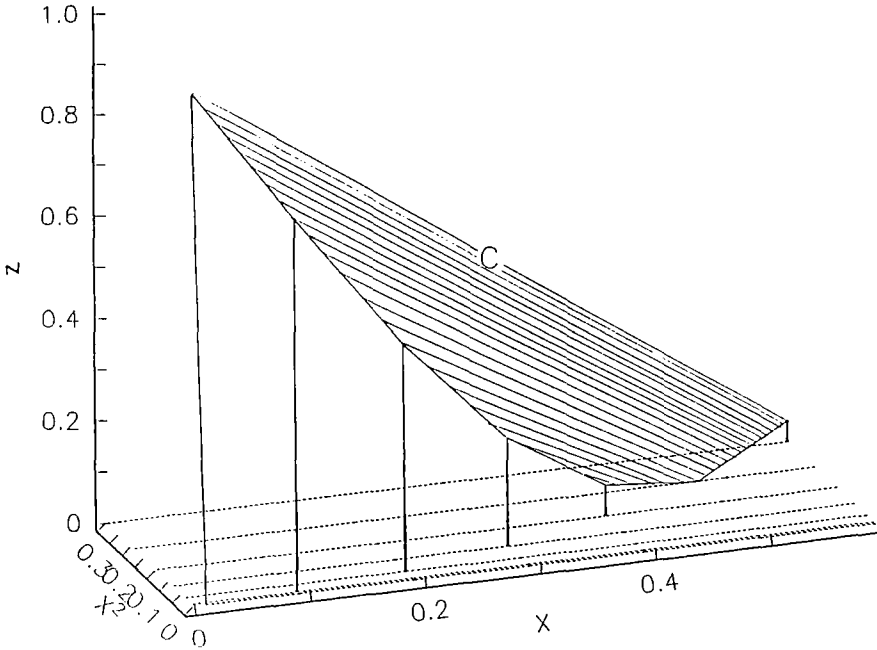


Fig. 1. The generalized moment problem in \mathbb{R}^1 .

true distribution is \bar{x} , and the second moment is $x^{(2)}$. The problem is illustrated geometrically in fig. 1. Here, $X = [0, 0.6]$, the vertical lines are $(i/10, (i/10)^2, t)$ for $0 \leq t \leq f(i/10)$, $i = 0, \dots, 6$, and C is the shaded convex hull of $(x, x^2, f(x))$ for $x \in X$. The objective in (5.1) is to find $y^* = (\bar{x}, x^{(2)}, z^*) \in C$ that maximizes z . A generalization of Carathéodory's theorem [39] for the convex hull of a connected set tells us that y^* can be expressed as a convex combination of at most three extreme points of C , giving us a special case of theorem 3.1. Therefore, an optimal solution to (5.1) can be written, $\{x^*, p^*\}$, where the points of support, $x^* = \{x_1^*, x_2^*, x_3^*\}$ have probabilities $p^* = \{p_1^*, p_2^*, p_3^*\}$. An optimal solution may, however, have two points of support. A function that has this property for a given instance of (5.1) is called a *two-point support* function. We give sufficient conditions below for a function to have this two-point support property. This property then allows a simplified solution of (5.1).

In this characterization, we use a dual problem to (5.1), the *semi-infinite* program [20] that appears, for example, in Chebyshev approximation [35]. For the one-dimensional, two-moment constraint problem considered here, this dual is to find θ, π_1, π_2 such that

$$\begin{aligned} \theta + \pi_1 x + \pi_2 x^2 &\geq f(x), \quad \forall x \in X, \\ \text{and } \theta + \pi_1 \bar{x} + \pi_2 x^{(2)} &\text{ is minimized.} \end{aligned} \tag{5.2}$$

Note that (5.2) involves three variables and an infinite number of constraints in contrast to the infinite dimensional, finitely constrained *primal problem* (5.1). Note also that an optimal solution to (5.2) is a quadratic function that dominates f and that has minimum expectation with respect to any probability measure in \mathcal{P} .

The optimality conditions on a feasible solution to (5.1), $x^* = \{x_1^*, x_2^*, x_3^*\}$, with associated probabilities, $p^* = \{p_1^*, p_2^*, p_3^*\}$, are that there exist dual variables, $\theta^*, \pi_1^*, \pi_2^*$, such that

$$\begin{aligned} \theta^* + \pi_1^* x_i^* + \pi_2^* (x_i^*)^2 &= f(x_i^*), & \text{if } p_i^* > 0, \\ \theta^* + \pi_1^* x + \pi_2^* x^2 &\geq f(x), & \forall x \in X, \end{aligned} \tag{5.3}$$

where the first condition is *complementary slackness* condition and the second condition is *dual feasibility*. A useful interpretation of these conditions in terms of the function ρ defined in (3.5) is that x_i^* has a positive probability, p_i^* , only if $\rho(x_i^*, \theta^*, \pi^*) = 0$ and x_i^* maximizes ρ over X for fixed (θ^*, π^*) . It is convenient to let $\rho(x, \theta, \pi) = f(x) - q(x, \theta, \pi)$, where $q(x, \theta, \pi) = \theta + \pi_1 x + \pi_2 x^2$. The following lemmata give additional properties for an optimal solution $\{x^*, p^*\}$ to (5.1). We assume that f is always convex below.

LEMMA 5.1

If $x = \{x_1, x_2, x_3\}$ is feasible in (5.1) with corresponding probabilities, $p = \{p_1, p_2, p_3\}$, $p_i > 0, i = 1, 2, 3$, then there exists another feasible solution $x' = \{x_1, x_4\}$, with $p' = \{p'_1, p_4\}$, where $x_2 < x_4 < x_3$.

Proof

Feasibility of $\{x_1, x_2, x_3\}$ and $\{p_1, p_2, p_3\} > 0$ implies that $\sum_{i=1}^3 p_i(x_i, (x_i)^2) = (\bar{x}, x^{(2)})$. Consider $t(x)$ such that $x_1(1 - t(x)) + t(x)x = \bar{x}$ and $w(x) = x_1^2(1 - t(x)) + (t(x))x^2$. Note that $t(x) = (\bar{x} - x_1)/(x - x_1)$ is in $(0, 1)$ for $x > \bar{x}$ and $w(x)$ is strictly increasing in x for $x > \bar{x}$ ($w'(x) = \bar{x} - x_1$). Now consider p and q such that $p(x_2 - x_1) + q(x_3 - x_1) + x_1 = \bar{x}$. Fixing p in this equation, we can find a corresponding $q(p) = [(\bar{x} - x_1) - p(x_2 - x_1)]/(x_3 - x_1)$. The weighted square value as a function of p and $q(p)$ is $u(p) = p(x_2^2 - x_1^2) + q(p)(x_3^2 - x_1^2) + x_1^2 = p(x_2 - x_1)(x_2 + x_1) + (\bar{x} - x_1)(x_3 + x_1) + x_1^2$. Since $t(x_2) > p_2 > 0$ and $x_3 > x_2 > x_1$, $w(x_2) = u(t(x_2)) < u(p_2) < u(0) = w(x_3)$. So, for w continuous and increasing, there exists x_4 such that $w(x_4) = u(p_2) = x^{(2)}$ and $x_2 < x_4 < x_3$. Letting $p' = \{1 - t(x_4), t(x_4)\}$ completes the proof. \square

The next lemma considers f with a derivative f' that has local convexity or concavity properties. These properties form the basis for the bounding approach given in this paper.

LEMMA 5.2

If $\rho(\hat{x}, \hat{\theta}, \hat{\pi}) = 0$ for some $(\hat{\theta}, \hat{\pi})$ feasible in (5.2), \hat{x} is in the interior of X , and f' is convex on $A = (\hat{x} - \eta, \hat{x})$ for some $\eta > 0$, then $f'(t) \geq q'(t, \hat{\theta}, \hat{\pi})$ for $t \in A$. Also, if f' is concave on $B = (\hat{x}, \hat{x} + \eta)$ for some $\eta > 0$, then $f'(t) \leq q'(t, \hat{\theta}, \hat{\pi})$ for $t \in B$.

Proof

First, we assume f' is convex on $A = (\hat{x} - \eta, \hat{x})$ for some $\eta > 0$. Let $\hat{\rho}(\cdot) = \rho(\cdot, \hat{\theta}, \hat{\pi})$ and let $\hat{q}(\cdot) = q(\cdot, \hat{\theta}, \hat{\pi})$. Note that \hat{q} is twice differentiable on X ($\hat{q}'(x) = \pi_1 + 2\pi_2x$, $\hat{q}''(x) = 2\pi_2$) and that, for f convex, f is differentiable on all but a countable number of points in X . Moreover, since f' is convex on A , f is also continuously differentiable on A . Hence, $\hat{\rho}'$ exists on A . We can let $\hat{\rho}'_-(\hat{x}) = \lim_{t \downarrow 0} \hat{\rho}'(\hat{x} - t)$. We can also define $\hat{\rho}'_+(\hat{x}) = \lim_{t \downarrow 0} \hat{\rho}'(\hat{x} + t)$ when f' is concave on B . Note that $\hat{\rho}'_{\pm}(\hat{x}) = f'_{\pm}(\hat{x}) - q'(\hat{x})$.

If $\hat{\rho}(\hat{x}) = 0$ for feasible $(\hat{\theta}, \hat{\pi})$ in (5.2), then \hat{x} maximizes $\hat{\rho}$ over X . Therefore, $0 \in \text{co}[\hat{\rho}'_-(\hat{x}), \hat{\rho}'_+(\hat{x})]$. By convexity of f , $f'_+(\hat{x}) \geq f'_-(\hat{x})$, $\hat{\rho}'_+(\hat{x}) \geq \hat{\rho}'_-(\hat{x})$, and, thus, $\hat{\rho}'_-(\hat{x}) \leq 0$. Suppose $\hat{\rho}'_-(\hat{x}) < 0$, then $\hat{\rho}(\hat{x} - t) + \hat{\rho}'(\hat{x} - s)(t) = \hat{\rho}(\hat{x})$ for all $t \in (0, \eta)$ and some $0 < s < t$. By $\hat{\rho}'$ continuous in this interval, there exists ϵ such that $\hat{\rho}'(\hat{x} - s) < 0$ for all $0 < s < \epsilon$. For $0 < t < \epsilon$, we would have $\hat{\rho}(\hat{x} - t) > \hat{\rho}(\hat{x})$, contradicting the maximality of \hat{x} . Hence, $\hat{\rho}'_-(\hat{x}) = 0$.

By applying the mean value theorem to $\hat{\rho}$ and $\hat{\rho}'$, we also have that $\hat{\rho}(\hat{x} - t) = \hat{\rho}(\hat{x}) + \hat{\rho}'(\hat{x} - s)(-t) = \hat{\rho}(\hat{x}) + (-t)(-s)(y(r) - \hat{q}''(r))$, where $0 < r < s < t$ and $y(r) \in [(f')'_-(\hat{x} - r), (f')'_+(\hat{x} - r)]$. Letting r vary with t as $r(t)$ and noting that $\lim_{t \rightarrow 0} y(r(t)) = (f')'_-(\hat{x})$ and $\hat{\rho}(\hat{x} - t) - \hat{\rho}(\hat{x}) \leq 0$ for t close to 0, we obtain $0 \geq (f')'_-(\hat{x}) - \hat{q}''(\hat{x})$. Now, consider $\hat{\rho}'(\hat{x} - t) = f'(\hat{x} - t) - \hat{q}'(\hat{x} - t) = f'(\hat{x}) - \hat{q}'(\hat{x}) + (-t)(y(\hat{x} - s) - \hat{q}''(\hat{x} - s))$ where $y(\hat{x} - s) \in [(f')'_-(\hat{x} - s), (f')'_+(\hat{x} - s)]$. By f' convex on A and $\hat{q}''(\hat{x} - s) = \hat{q}''(\hat{x}) = \pi_2$ for any s , $y(\hat{x} - s) - \hat{q}''(\hat{x} - s) \leq (f')'_-(\hat{x}) - \hat{q}''(\hat{x}) \leq 0$. So, $f'(\hat{x} - t) - \hat{q}'(\hat{x} - t) \geq 0$ for all $t \in (0, \eta)$. A similar argument holds if f is concave on B . \square

The previous lemma considers local convexity properties of f' when it exists. The following results refer to functions with derivatives that are convex and then concave.

LEMMA 5.3

Let $g(x) = h(x) - c(x)$ be a function such that $h(x)$ is increasing and upper semi-continuous on \mathbb{R} , $h(x)$ is convex on $(-\infty, y)$ and concave on (y, ∞) , and $c(x)$ is an affine function on \mathbb{R} . Then there exists a partition of $(-\infty, \infty)$ into subintervals, $I_1 = (-\infty, a_1]$, $I_2 = (a_1, a_2)$, $I_3 = [a_2, a_3)$, $I_4 = [a_3, +\infty)$, $-\infty \leq a_1 \leq a_2 \leq a_3 \leq \infty$, such that $g(x) \geq 0$ for all $x \in I_1 \cup I_3$ and $g(x) \leq 0$ for all $x \in I_2 \cup I_4$. (When any of $a_1, a_2, a_3 = \pm\infty$, then we interpret the interval as open at the corresponding value of $\pm\infty$.)

Proof

First note that g is continuous on $(-\infty, y)$ and (y, ∞) . By convexity and for h increasing at y , $S_0^- = \{x \mid g(x) \leq 0, x < y\} = [a, b]$ (or \emptyset) for some $-\infty \leq a \leq b \leq y$ (when $a = -\infty$ and when $b = y$, the interval is open at a or b respectively). Similarly, by concavity on (y, ∞) , $S_0^+ = \{x \mid g(x) \geq 0, x \geq y\} = [c, d]$ (or \emptyset) for some $\infty \geq d \geq c \geq y$ (where we have an open interval if $d = \infty$). Note that if $b < y$, then $g(x) \geq 0$ on $[b, y]$, so $c = y$. From this observation, if S_0^- and S_0^+ are not empty, then g is nonnegative on $(-\infty, a]$ and either $[b, d]$ or $[c, d]$ by continuity at a and b if $b < y$. If $S_0^- = \emptyset$, then $S_0^+ \neq \emptyset$, and g is nonnegative on $(-\infty, d]$, since g is increasing at y . If $S_0^+ = \emptyset$, then g is nonnegative on $(-\infty, a]$. Thus, let $a_1 = a$ if $S_0^- \neq \emptyset$; $a_1 = d$ if $S_0^- = \emptyset$, $a_2 = b$ if $S_0^- \neq \emptyset$ and $b < y$, $a_2 = c$ if $S_0^- \neq \emptyset$ and $b = y$, $a_2 = \infty$ if $S_0^- = \emptyset$, and $a_3 = d$ if $S_0^- \neq \emptyset$, $S_0^+ \neq \emptyset$, $a_3 = \infty$ if $S_0^- = \emptyset$ or $S_0^+ = \emptyset$. This yields the regions in the lemma. \square

The next lemma considers the case where $g = \rho'$ is constant on an interval. In the following, we use the notation $X = [a, b]$ for convenience. It is assumed that this also includes the cases $X = (-\infty, b]$, $X = [a, +\infty)$, and $X = (-\infty, +\infty)$ unless explicitly stated otherwise.

LEMMA 5.4

If f is convex on $X = [a, b]$ with derivative f' defined as a convex function on $[a, c)$ and as a concave function on $(c, b]$ for $a \leq c \leq b$ and if $\rho(x, \hat{\theta}, \hat{\pi}) = 0$ for some $(\hat{\theta}, \hat{\pi})$ feasible in (5.2) and for all $x \in (\hat{x} - \epsilon, \hat{x} + \epsilon)$ for some $\hat{x} \in X$ and $\epsilon > 0$, then there exists an interval $D \supset (\hat{x} - \epsilon, \hat{x} + \epsilon)$ such that $\rho(x, \hat{\theta}, \hat{\pi}) = 0$ for all $x \in D$ and $\rho(x, \hat{\theta}, \hat{\pi}) < 0$ for all $x \notin$ the closure of D .

Proof

Let $D = (d, e)$ be the largest open interval including $(\hat{x} - \epsilon, \hat{x} + \epsilon)$ such that $\hat{\rho}(x) = 0$ for all $x \in D$. First, we show that $c \in [d, e]$. If not, then suppose $d > c$. In this case, f' is concave on $(c, e]$. For $x \in D$, $\hat{\rho}(x) = 0$, so $f'(x) = \hat{q}'(x)$ and $f''(x) = \hat{q}''(x)$. From lemma 5.1, $\hat{q}'(x) \geq f'(z)$ on (c, e) , and $f'(x) - \hat{q}'(x) < 0$ on (c, d) by concavity. Thus, $\hat{\rho}(\hat{x}) = \hat{\rho}(d) < \hat{\rho}(x)$ for any $x \in (c, d)$, which contradicts the maximality of \hat{x} . Hence, $d \leq c$. Symmetrically, $e \geq c$.

By lemma 5.2 and f' 's implied semicontinuity at the endpoints, $\hat{\rho}' \geq 0$ on $[a, c)$ and $\hat{\rho}' \leq 0$ on $(c, d]$. By convexity of f' on $[a, c)$, the set $\{x \mid f'(x) \leq \hat{q}'(x), a \leq x < c\} = [d, c)$, if $d < a$. Hence, $\hat{\rho}'(x) > 0$ on $[a, d)$. Similarly, $\hat{\rho}'(x) < 0$ on $(e, b]$. Thus, $\hat{\rho}(x) < 0$ on $[a, d)$ and $(e, b]$. \square

The convex-concave property is now used to derive our main result about two-point support functions.

THEOREM 5.1

If f is convex with derivative f' defined as a convex function on (a, c) and as a concave function on (c, b) for $X = [a, b]$ and $a \leq c \leq b$, then there exists an

optimal solution to (5.2) with at most two support points, $\{x_1, x_2\}$, with positive probabilities, $\{p_1, p_2\}$.

Proof

Let $\{\hat{\theta}, \hat{\pi}\}$ be an optimal solution to (5.2). First, assume that there does not exist $\epsilon > 0$, $\hat{x} \in (a, b)$ such that $\rho(x, \hat{\theta}, \hat{\pi}) = 0$ for all $x \in (\hat{x} - \epsilon, \hat{x} + \epsilon)$. By lemmas 5.1 and 5.3, the only isolated points where $\hat{\rho}$ could be 0 and maximized are a_1 and a_3 if $[a, b] \supset I_2 \cup I_3$. If $[a, b] \not\supset I_2$, then a can replace a_1 and if $[a, b] \not\supset I_3$, then b can replace a_3 , but, in either case, at most two points meet the conditions for optimality.

If there exists $\epsilon > 0$, $\hat{x} \in (a, b)$ such that $\rho(x, \hat{\theta}, \hat{\pi}) = 0$ for all $x \in (\hat{x} - \epsilon, \hat{x} + \epsilon)$, then lemma 5.4 implies that any optimal solution $\{x_1, x_2, x_3\}$ must be in the closure of D and that $\hat{\rho}(x) = 0$ for all $x \in D$. By lemma 5.2, we can select x_4 in (x_2, x_3) such that there exists $\{p_1, p_4\}$ so that $\{x_1, x_4, p_1, p_4\}$ is feasible in (5.1). The optimality conditions still hold for $\hat{\rho}(x_4) = 0$. Hence, $\{x_1, x_4, p_1, p_4\}$ is optimal in (5.1). \square

A corollary of theorem 5.1 is that any function f that has a convex or concave derivative has the two-point support property. The class of functions that meets the criteria of theorem 5.1 contains many useful examples. Some of these functions are given below:

- (1) Polynomials defined over ranges with at most one third derivative sign change.
- (2) Exponential functions of the form, $c_0 e^{c_1 x}$, $c_0 \geq 0$.
- (3) Logarithmic functions of the form, $\log_k(cx)$, for any $k \geq 0$.
- (4) Certain hyperbolic functions such as $\sinh(cx)$, $c, x \geq 0$, $\cosh(cx)$.
- (5) Certain trigonometric and inverse trigonometric functions such as $\tan^{-1}(cx)$, $c, x \geq 0$.

In fact, theorem 5.1 can be applied to provide an upper bound on the expectation of any convex function with known third derivative when the distribution function has a known third moment, $x^{(3)}$. Suppose $a > 0$ (if not, then this argument can be applied on $[a, 0]$ and $[0, b]$), then let $g(x) = \alpha x^3 + f(x)$. The function g is still convex on $[0, b]$ for $\alpha \geq 0$. By defining $\alpha \geq (-1/6)\min(0, \inf_{x \in [a, b]} f'''(x))$, g' is convex on $[a, b]$, and an upper bound, $UB(g)$, on $E_g(x)$ has a two-point support. The expectation of f is then bounded by

$$E_g(x) \leq UB(g) - \alpha x^{(3)}. \tag{5.4}$$

The conditions in theorem 5.1 are only sufficient for a two-point support function. They are not necessary. The following function, for example, has an optimal two-point support at $x^* = \{1/3, 1\}$ for any corresponding feasible p_1

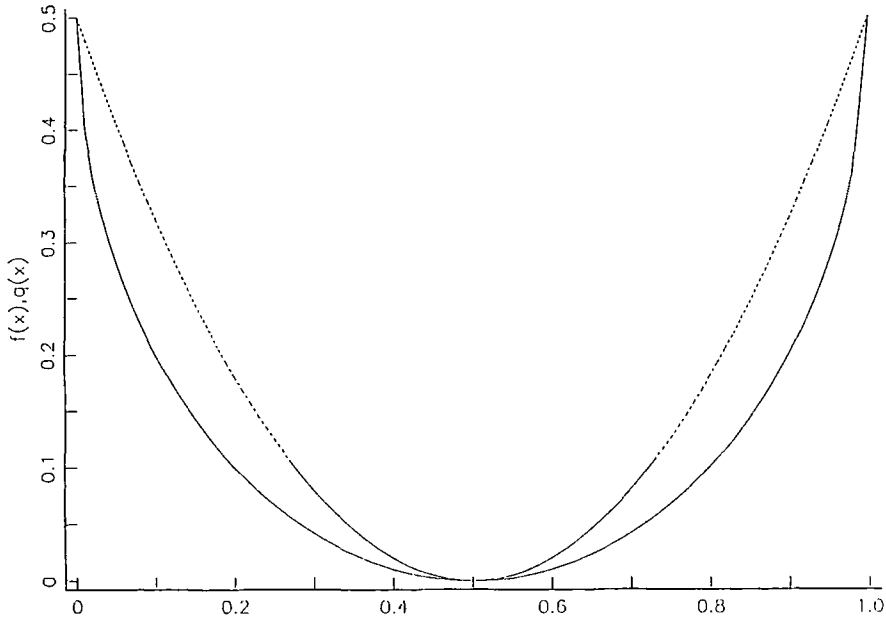


Fig. 2. A function requiring three support points.

and p_2 when $X = [0, 1]$.

$$f(x) = \begin{cases} 6/5 - 4x + 5x^2 & \text{if } 0 \leq x < 0.2, \\ 2x + 1 & \text{if } 0.2 \leq x < 0.4, \\ -(2/5)x - 8x^2 + 10x^3 & \text{if } 0.4 \leq x < 0.6, \\ -4x + 4x^2 & \text{if } 0.6 \leq x \leq 1. \end{cases} \tag{5.5}$$

The function defined in (5.5) does not, however, meet the conditions of theorem 5.1. This and other two-point support functions can always be constructed by fixing θ , π_1 , π_2 and considering any function $f(x) \leq q(x, \theta, \pi)$ such that $f(x_1) = q(x_1, \theta, \pi)$, $f(x_2) = q(x_2, \theta, \pi)$, $\lambda x_1 + (1 - \lambda)x_2 = \bar{x}$, and $\lambda x_1^2 + (1 - \lambda)x_2^2 = x^{(2)}$ for some $0 < \lambda < 1$. Any convex f satisfying these conditions corresponds to θ , π_1 , π_2 optimal in (5.2). No other conditions on the function are necessary.

Note also that not all functions are two-point support functions (although bounds using (5.4) are available). A function requiring three support points, for example, is $f(x) = (1/2) - \sqrt{(1/4) - (x - (1/2))^2}$. This function and its optimal dominating quadratic function are illustrated in fig. 2.

Given that a function is a two-point support function, the points $\{x_1, x_2\}$ can be found using a line search to find a maximum. For example, if some candidate

Table 1
Bounding values

Function	\bar{x} ,	$x^{(2)}$	Jensen	Beta	2-M	S-L	E-M
e^{-x}	0.500,	0.333	0.607	0.622	0.624	0.651	0.684
x^3	0.833,	0.714	0.579	0.625	0.629	0.675	0.833
$\sin(\pi(x+1))+1$	0.500,	0.333	0.000	0.363	0.384	0.577	1.000

$x_1 < \bar{x}$ is given, then a feasible corresponding x_2 is

$$x_2 = \frac{x^{(2)} - \bar{x}x_1}{\bar{x} - x_1}, \tag{5.6}$$

where $p_1 = (x_2 - \bar{x}) / (x_2 - x_1)$ and $p_2 = 1 - p_1$. Note that the problem is obviously not feasible if $(x_1, x_2) \not\in \bar{x}$. The solution of (5.1) then reduces to maximizing:

$$\gamma(x_1) = p_1(x_1)f(x_1) + p_2(x_1)f(x_2(x_1)) \tag{5.7}$$

subject to $x_1 \in [a, \bar{x}]$.

A line search to find the maximum in (4.7) can be performed efficiently using, for example, Lemaréchal and Mifflin's procedure in [30] if $\gamma \in C^2$ or Mifflin and Strodriot's [32] method without derivatives. Table 1 gives the values (under "2-M") that were obtained by this procedure for three two-point support functions with distributions on $[0, 1]$. Figures 3–5 illustrate that the optimal points, $\{x_1^*, x_2^*\}$, may be at either endpoint or interior to $[0, 1]$ and how the bounding

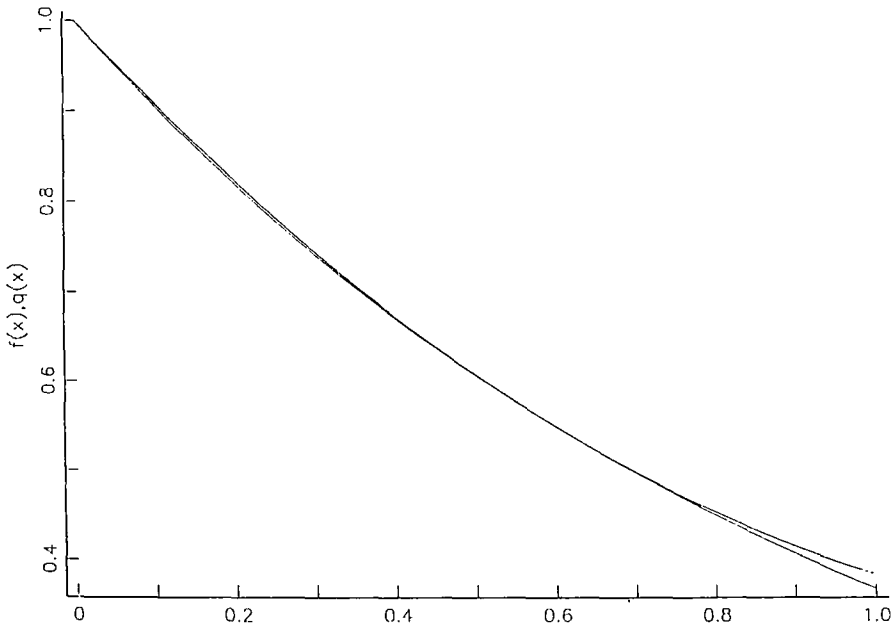


Fig. 3. Optimal bounding function for e^{-x} .

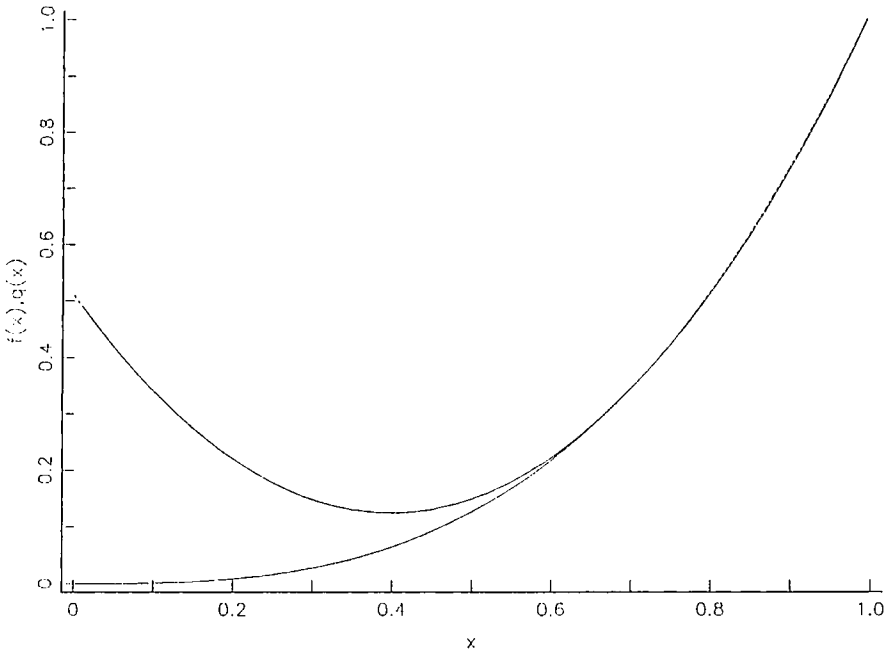


Fig. 4. Optimal bounding function for x^3 .

functions vary. The table gives the Jensen lower bound and the expectation for a random variable with beta distribution (under “Beta”) with the given first and second moments for comparison. The Edmundson–Madansky upper bound (“E–M”) is also provided. The “S–L” value (for “semi-linear bound”) given in table 1 corresponds to bounding f with a *semi-linear* function that has the form:

$$f(x) = \begin{cases} q^-(c - x) & \text{if } x \leq c, \\ q^+(x - c) & \text{if } x > c, \end{cases}$$

where $q^+ + q^- \geq 0$. This type of function is useful because a line search is not necessary for solving (5.7). The support points can be calculated analytically by observing that the conditions of theorem 5.1 are met and by finding the optimal x_1 in (5.7) as a function of c . The results depend on the interval, $[a, b]$. If $[a, b] = [0, 1]$, then consider the nonintersecting intervals, $A = (0, x^{(2)}/(2\bar{x}))$, $B = [x^{(2)}/(2\bar{x}), (1 - x^{(2)})/(2(1 - \bar{x}))]$, and $C = ((1 - x^{(2)})/(2(1 - \bar{x})), 1)$. The points of support for a semi-linear, convex function defined on $[0, 1]$ are

$$\{x_1^*, x_2^*\} = \begin{cases} \{0, x^{(2)}/\bar{x}\} & \text{if } c \in A, \\ \{c - d, c + d\} & \text{if } c \in B, \\ \{(\bar{x} - x^{(2)})/(1 - \bar{x}), 1\} & \text{if } c \in C, \end{cases} \tag{5.8}$$

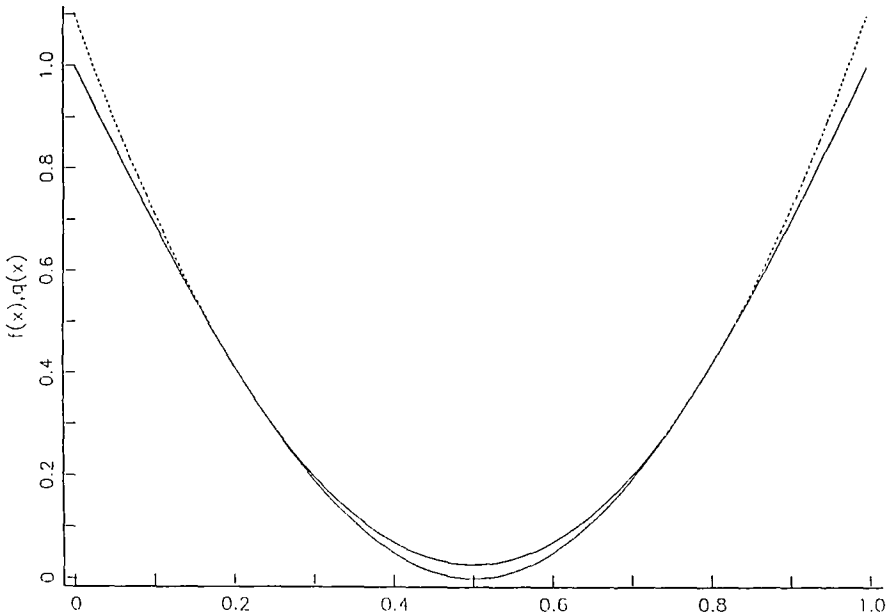


Fig. 5. Optimal bounding function for $\sin(\pi(x + 1)) + 1$.

where $d = \sqrt{c^2 + 2c\bar{x} + x^{(2)}}$. This result can be obviously extended to all finite intervals. It results from using (5.6) and differentiating γ with respect to x_1 in (5.7).

Infinite intervals can also be solved analytically for semi-linear, convex functions. For $X = [0, \infty)$, the results are as in (5.8) with $B = [x^{(2)}/(2\bar{x}), \infty)$ and $C = \emptyset$. For the interval $(-\infty, \infty)$, the points of support are those for interval B in (5.8). We note that special cases for these supports of semi-linear, convex functions were considered in [11,22,38].

Semi-linear, convex functions are common in decision problems to represent penalties for being above or below a preferred value, c . They can also be used, however, to provide bounds for other convex functions when only the first and second moments of the distribution function are known. Results from using these functions in problems with linear recourse appear in [2,3,40].

We conclude with an example for bounding a nonlinear recourse function with the form in (1.3). We suppose in this case that

$$f(x(1), x(2)) = \begin{cases} \min_{y_1, y_2} (y_1 - 1)^2 + (y_2 - 2)^2 \\ \text{s.t. } y_1^2 + y_2^2 - 1 \leq x(1), \\ (y_1 - 1)^2 + y_2^2 - 1 \leq x(2). \end{cases} \tag{5.9}$$

This problem may correspond to determining a performance characteristic of a part that is machined by two circular motions centered at $(0, 0)$ and $(1, 0)$ respectively. Here, the performance characteristic is proportional to the distance from the finished part to another object at $(2, 1)$. The square of the radii of the tool motions is $x(i) + 1$ where $x(i)$ is a nonnegative random variable associated with the machines' precision. Figure 6 illustrates the solutions for $x(i) = 0$ and $x(i) = 1$ for $i = 1, 2$.

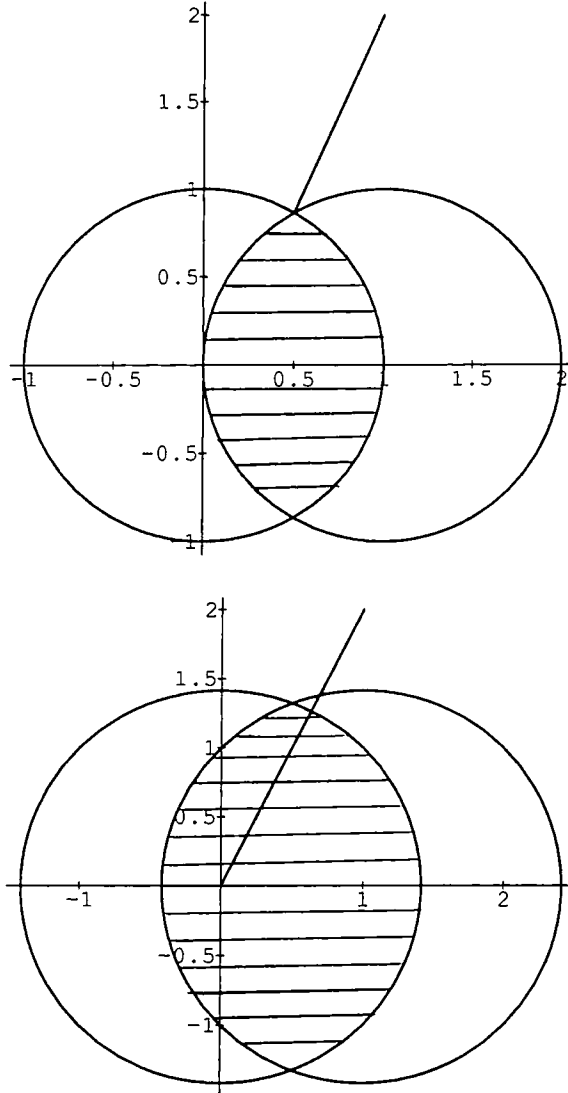


Fig. 6. Nearest points to $(2, 1)$.

We suppose that it is known that the $x(i)$ are nonnegative, that $\bar{x}(i) = 1$ and the $x^{(2)}(i) = 1.25$. In this case, we would like an upper bound on the expected performance $E_f(x)$. We construct a bound by first finding $\nu_i(x(i))$ as in (4.1). In this case, we solve (5.9) with right-hand sides $mx(i)e_i$ to obtain

$$\nu_1(x(1)) = \frac{1}{2} \begin{cases} \frac{(2x(1) - 1)^2}{4} + \left(\frac{(3 + 4x(1) - 4x(1)^2)^{0.5}}{2} - 2 \right)^2 & \text{if } x(1) \leq 0.5, \\ 1.0 & \text{if } x(1) > 0.5; \end{cases}$$

$$\nu_2(x(2)) = \frac{1}{2} \begin{cases} \frac{(2x(2) + 1)^2}{(\sqrt{5} - 1)^2} + \left(\frac{(3 + 4x(2) - 4x(2)^2)^{0.5}}{2} - 2 \right)^2 & \text{if } x(2) \leq 0.5 - \left(\frac{\sqrt{5}}{5} \right), \\ (\sqrt{5} - 1)^2 & \text{if } x(2) > 0.5 - \left(\frac{\sqrt{5}}{5} \right). \end{cases}$$

These functions are illustrated in fig. 7. Their third derivatives are nonpositive so the first derivative of each ν_i is concave, meeting the conditions for two-point

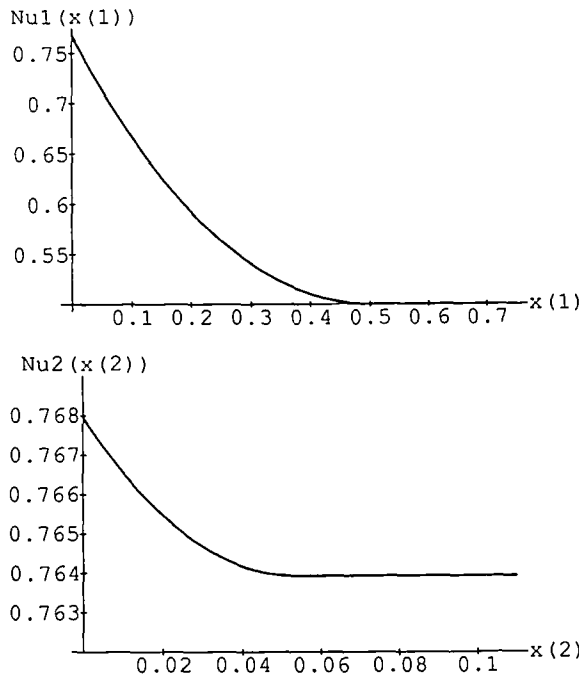


Fig. 7. The bounding functions ν_1 and ν_2 .

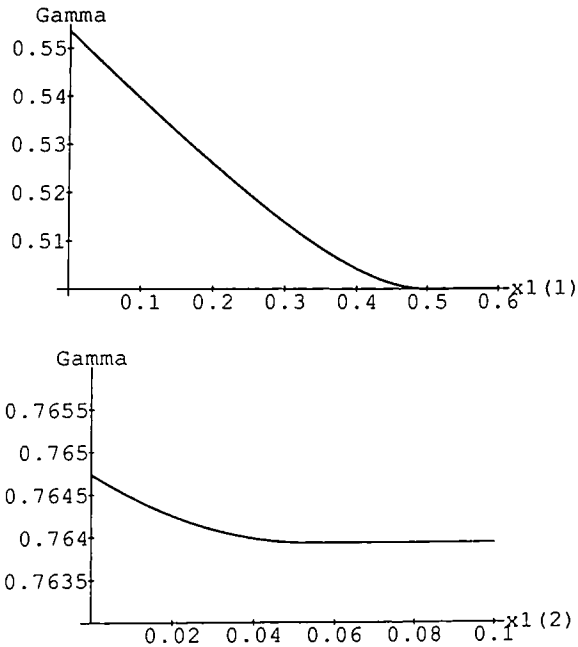


Fig. 8. The γ functions for $x(1)$ and $x(2)$.

support. The functions γ for $x(1)$ and $x(2)$ are given in fig. 8. Notice that the maximum occurs at 0 in each case. The resulting upper bound on $E_f(x)$ is then $(1.107 + 1.529)/2 = 1.318$. We can compare this to the lower bound of $f(\bar{x}) = 0.822$ to bound the expected performance. In comparing to other bounds, the best available bound knowing only the mean and the nonnegative range is 1.529, the value at $(0, 0)$, since a distribution can meet the range and mean conditions with a probability p mass at $(0, 0)$, probability $(1-p)q$ at $(\bar{x}(1), 0)/q(1-p)$, and probability $(1-p)(1-q)$ at $(0, \bar{x}(2))/(1-p)(1-q)$, where $0 < q < 1$ and p may become arbitrarily close to 1.

6. Conclusions

This paper describes a bound on the general recourse problem and a resulting bound on the expectation of convex functions when only limited distributional information is available. Given the difficulties in estimating random phenomena, limited information in terms of bounds on the mean and second moments of distributions is a general practical situation. The bounds provided in this paper allow for efficient computation. We note that these procedures can be extended to lower bounds on the expectation of concave functions obviously. Since the

Jensen inequality is the solution of the generalized moment problem to minimize the expectation of a convex function subject to a first moment equality constraint and an upper bound on the second moment constraint, upper and lower bounds are computable on the expectation of general functions that can be expressed as linear combinations of convex and concave functions. Given this extension, the use of separable convex functions and the two-point support bounds apply to a wide range of problems.

References

- [1] A. Ben-Tal and E. Hochman, More bounds on the expectation of a convex function of a random variable, *J. Appl. Prob.* 9 (1972) 803–812.
- [2] J.R. Birge and S.W. Wallace, A separable piecewise linear upper bound for stochastic linear programs, *SIAM J. Control Optim.* 26 (1988) 725–739.
- [3] J.R. Birge and R.J-B. Wets, Designing approximation schemes for stochastic optimization problems, in particular, for stochastic programs with recourse, *Math. Progr. Study* 27 (1986) 54–102.
- [4] J.R. Birge and R.J-B. Wets, Sublinear upper bounds for stochastic linear programs with recourse, *Math. Progr.* 43 (1989) 131–149.
- [5] J.R. Birge and R.J-B. Wets, Computing bounds for stochastic programming problems by means of a generalized moment problem, *Math. Oper. Res.* 12 (1987) 149–162.
- [6] T. Cipra, Moment problem with given covariance structure in stochastic programming, *Ekonom.-Mat. Obzor* 21 (1985) 66–77.
- [7] G.B. Dantzig, *Linear Programming and Extensions* (Princeton University Press, Princeton, NJ, 1963).
- [8] P.J. Davis and P. Rabinowitz, *Methods of Numerical Integration* (Academic Press, Orlando, FL, 1984).
- [9] J. Dulá, Bounds on the expectation of convex functions, Ph.D. dissertation, The University of Michigan, Ann Arbor, Michigan (1986).
- [10] J. Dulá, An upper bound on the expectation of simplicial functions of multivariate random variables, *Math. Progr.*, to appear.
- [11] J. Dupačová, Minimax stochastic programs with nonconvex nonseparable penalty functions, in: *Progress in Operations Research*, A. Prékopa (ed.) (North-Holland, Amsterdam, 1976) pp. 303–316.
- [12] J. Dupačová, Minimax approach to stochastic linear programming and the moment problem. Recent results, *ZAMM* 58 (1977) T466–467.
- [13] H.P. Edmundson, Bounds on the expectation of a convex function of a random variable, The Rand Corporation Paper 982, Santa Monica, CA (1956).
- [14] Y. Ermoliev, A. Gaivoronski and C. Nedeeva, Stochastic optimization problems with partially known distribution functions, *SIAM J. Control Optim.* 23 (1985) 696–716.
- [15] K. Frauendorfer, Solving SLP recourse problems with arbitrary multivariate distributions – the dependent case, *Math. Oper. Res.* 13 (1988) 377–394.
- [16] K. Frauendorfer, Solution methods for linear–quadratic stochastic optimization problems, presented at the *5th Int. Conf. on Stochastic Programming*, Ann Arbor, Michigan (August 1989).
- [17] K. Frauendorfer and P. Kall, A solution method for SLP recourse problems with arbitrary multivariate distributions – the independent case, *Problems Control Inf. Theory* 17 (1988) 177–205.

- [18] H. Gassmann and W.T. Ziemba, A tight upper bound for the expectation of a convex function of a multivariate random variable, *Math. Progr. Study* 27 (1986) 39–53.
- [19] D. Hausch and W.T. Ziemba, Bounds on the value of information in uncertain decision problems II, *Stochastics* 10 (1983) 181–217.
- [20] R. Hettich, An implementation of a discretization method for semi-infinite programming, *Math. Progr.* 34 (1986) 354–361.
- [21] C.C. Huang, W. Ziemba and A. Ben-Tal, Bounds on the expectation of a convex function of a random variable: with applications to stochastic programming, *Oper. Res.* 25 (1979) 315–325.
- [22] R. Jagannathan, A minimax procedure for a class of linear programs under uncertainty, *Oper. Res.* 25 (1977) 173–177.
- [23] J.L. Jensen, Sur les fonctions convexes et les inégalités entre les valeurs moyennes, *Acta Math.* 30 (1906) 175–193.
- [24] P. Kall, Stochastic programming with recourse: upper bounds and moment problems – a review, in *Advances in Mathematical Optimization*, J. Guddat (ed.) (Akademie-Verlag, Berlin, 1988) pp. 86–103.
- [25] P. Kall, An upper bound for SLP using first and total second moments, this volume.
- [26] P. Kall and D. Stoyan, Solving stochastic programming problems with recourse including error bounds, *Math. Operationsforsch. Statist. Ser. Optim.* 13 (1982) 431–447.
- [27] A. Karr, Extreme points of certain sets of probability measures. with applications, *Math. Oper. Res.* 8 (1983) 74–85.
- [28] J.H.B. Kemperman, The general moment problem, a geometric approach, *Ann. Math. Statist.* 39 (1968) 93–122.
- [29] M. Krein and A. Nudelman, The Markov moment problem and extremal problems. Translation of *Mathematical Monographs*, 50 (The American Mathematical Society, Providence, RI, 1977).
- [30] C. Lemaréchal and R. Mifflin, Global and superlinear convergence of an algorithm for one-dimensional minimization of convex functions, *Math. Progr.* 24 (1982) 241–256.
- [31] A. Madansky, Bounds on the expectation of a convex function of a multivariate random variable, *Ann. Math. Statist.* 30 (1959) 743–746.
- [32] R. Mifflin and J. Strodios, Reliable and rapid univariate minimization without derivatives, presented at *TIMS/ORSA Joint National Meeting*, Los Angeles, CA (April, 1986).
- [33] A.C. Miller, III, and T.R. Rice, Discrete approximations of probability distributions, *Manag. Sci.* 29 (1983) 352–362.
- [34] G. Peano, Residuo in formulas de quadratura, *Mathesis* 39 (1914) 5–10.
- [35] T.J. Rivlin, *An Introduction to the Approximation of Functions* (Blaisdell, Waltham, MA, 1969).
- [36] R.T. Rockafellar, *Convex Analysis* (Princeton University Press, Princeton, NJ, 1970).
- [37] N.M. Roy, Extreme points of convex sets in infinite dimensional spaces, *Am. Math. Monthly* 94 (1987) 409–422.
- [38] H. Scarf, A minimax solution of an inventory problem, in: *Studies in the Mathematical Theory of Inventory and Production*, K.J. Arrow, S. Karlin and H. Scarf (eds.) (Stanford University Press, Stanford, CA, 1958).
- [39] F.A. Valentine, *Convex Sets* (McGraw-Hill, New York, 1964).
- [40] S. Wallace, A piecewise linear upper bound on the network recourse problem, *Math. Progr.* 38 (1987) 133–146.
- [41] J. Žáčková, On minimax solutions of stochastic linear programming problems, *Časopis pro Pěstování Matematiky* 91 (1966) 423–430.