Abstract: Motivated by problems associated with the emerging theory of distributionally robust Monte Carlo simulation, this paper addresses the classical equation $Ax = b$ with $n \times n$ matrix $A = A(\theta)$ and $n \times 1$ vector $b = b(\theta)$ depending on an $m$-tuple of parameters $\theta$ with components $\theta_i$ entering in a rank-one manner. For such a system, the following convexity problem is considered: Determine if the second partial derivative of a solution component $x_i(\theta)$ with respect to a specified parameter $\theta_j$ is positive for all $\theta$ in a prescribed hypercube $\Theta_r$ of radius $r \geq 0$. The main result of this paper is an extreme point solution of this problem. To this end, a factorization of the second derivative of $x_i(\theta)$ is provided, which plays a major role in obtaining the so-called radius of convexity.

Keywords: Convex optimization, linear equations, probabilistic models, robustness, stochastic systems, uncertain linear systems

1. INTRODUCTION

This paper addresses a fundamental problem motivated by the emerging theory of distributionally robust Monte Carlo simulation. Preliminary results motivating this work are provided in (Ganesan et al., 2001) and basic references on distributional robustness include (Barmish and Lagoa, 1997), (Lagoa, 1998) and (Lagoa and Barmish, 2001). In this theory, the satisfaction of certain uniform convexity conditions is seen to facilitate computation; e.g., see (Barmish and Shcherbakov, 1999) and (Barmish, 2000). Whereas the use of traditional Monte Carlo simulation software requires probability distributions for the uncertain parameters, for example, see (Rubinstein, 1981), distributionally robust Monte Carlo simulation does not. In the distributional robustness setting, similar to classical robustness theory, the uncertain parameters are described solely in terms of their bounds with no a priori statistics assumed. Instead of conducting simulations using some rather arbitrary probability distribution such as uniform or Gaussian, robustness considerations lead to an entirely new approach. In a sense, this new approach to simulation addresses the robustician’s main objection to the classical Monte Carlo approach. Namely, if the statistics of the parameters are unknown and one simply imposes “reasonable” probability distributions for the sake of simulation, the resulting assessment of performance may turn out to be unrealistic.

1.1 Basic Problem Overview

To overview the basic problem considered in this paper, following (Ganesan et al, 2001), we consider the set of linear equations

$$Ax = b$$

with $n \times n$ matrix $A = A(\theta)$ and $n \times 1$ vector $b = b(\theta)$ depending on the $m$-tuple of parameters

$$\theta \doteq (\theta_1, \theta_2, \ldots, \theta_m).$$
in a hypercube $\Theta_\gamma$ of radius $r \geq 0$. The objective is to determine the largest radius $r$ guaranteeing that the $i$-th solution component $x_i(\theta)$ of

$$x(\theta) = A^{-1}(\theta)b(\theta)$$

not only exists but is either convex or concave with respect to individual components $\theta_j$. The main result applies when the parameter of interest $\theta_j$ enters into either $A(\theta)$ in a rank-one manner or $b(\theta)$ in a linear manner. That is, this dependence is said to be rank-one if there exist non-zero vectors $d^i, e^j \in \mathbb{R}^n$ such that

$$A(\theta) \doteq A_0 + \theta_j d^i (e^j)^T$$

with $A_0$ depending on $\theta_i$ for $i \neq j$. When $\theta_j$ enters into $b(\theta)$, the dependence is assumed to be affine linear. That is, there exists a non-zero vector $b^j \in \mathbb{R}^n$ such that

$$b(\theta) = b^0 + \theta_j b^j$$

with $b^0$ depending on $\theta_i$ for $i \neq j$. For such cases, these linear equations are said to have rank-one uncertainty structure. A more detailed problem formulation will be provided in Section 1.4 after the role of convexity on distributional robustness theory is explained.

### 1.2 Role of Convexity

To motivate the convexity issue considered in this paper, nominal $\theta = \theta^0$ is taken to be the known mean for an $m$-tuple of parameters $\theta$; these parameters are assumed to be independent with unknown joint probability distribution supported in a prescribed hypercube $\Theta_\gamma$ of radius $r \geq 0$. Now, a marginal probability density function $f_i(\theta_i)$ is said to be admissible for $\theta_i$ if it is symmetric and non-increasing with respect to $|\theta_i - \theta_i^0|$. The Dirac delta function at $\theta_i = \theta_i^0$ is also included as admissible and we denote the resulting family of joint probability density functions for $\theta$ by $\mathcal{F}$. With $f \in \mathcal{F}$ as the joint probability density function for $\theta$, we associate the random $m$-tuple $\theta^f$.

Now, with notation as above and $x_i(\theta)$ being some solution component of interest, existing theory guarantees that a distributionally robust expected value is readily obtainable under a convexity hypothesis; see (Lagoa and Barmish, 2001), (Barmish and Shcherbakov, 1999), and (Barmish, 2000). That is, with $u$ and $\delta$ denoting the uniform and Dirac delta function probability distributions respectively, it is known that if $x_i(\theta)$ is convex with respect to $\theta_j$ for $\theta \in \Theta_\gamma$, then the distributionally robust expected value

$$E^* \doteq \max_{f \in \mathcal{F}} E[x_i(\theta^f)]$$

is attained with probability density function $f^* \in \mathcal{F}$ having $j$-th marginal component $f^*_j = u$. Similarly, if $x_i(\theta)$ is concave with respect to $\theta_j$ for $\theta \in \Theta_\gamma$, then the distributionally robust expected value is attained with $f^*_j = \delta$.

### 1.3 Motivating Example

To provide a specific example addressed by the results in this paper, in the circuit of Figure 1.3, we make the identification between circuit variables and linear algebra variables used in this paper. The $\theta_i$ parameters of interest are taken to be the

![Fig. 1. Example circuit](image-url)

resistors, i.e., $\theta_1 \doteq R_1$, $\theta_2 \doteq R_2$, $\theta_3 \doteq R_3$, and $\theta_4 \doteq R_4$, and the solution variables $x_i$ are taken to be the currents and output voltage; i.e., $x_1 \doteq I_1$, $x_2 \doteq I_2$, $x_3 \doteq I_3$, $x_4 \doteq I_4$, and $x_5 \doteq V_{out}$. The theory in this paper is readily applied to obtain a distributionally robust Monte Carlo simulation of the expected value of the output voltage

$$x_5(\theta) = \frac{N(\theta)}{D(\theta)}$$

where the numerator and denominator, $N(\theta)$ and $D(\theta)$, are given by

$$N(\theta) = -200(100\theta_2 + 10\theta_3 + 10\theta_4 + \theta_1\theta_4) - 350(\theta_1\theta_4 + \theta_2\theta_4 + 2\theta_3\theta_4 + \theta_1\theta_2 + \theta_1\theta_3) - 30(\theta_2\theta_4 + \theta_1\theta_2\theta_4);$$

$$D(\theta) = 50(2\theta_1\theta_4 + 10\theta_4 + \theta_3\theta_4 - 10\theta_3 - \theta_2\theta_3) - 50(\theta_2\theta_1 - \theta_1\theta_3).$$

### 1.4 Precise Formulation

We let the nominal setting for the center of the parameter hypercube $\Theta_\gamma$ of radius $r \geq 0$ be denoted by

$$\theta^0 = (\theta_1^0, \theta_2^0, \ldots, \theta_m^0)$$

and consider the case when the solution $x(\theta^0)$ exists with second partial derivative of solution component $x_i(\theta)$ with respect to $\theta_j$ being positive. Accordingly, we assume

$$x^0 = A^{-1}(\theta^0)b(\theta^0)$$
exists and that the second partial derivative
\[ \nabla^2_{ij}(\theta) \doteq \frac{\partial^2 x_i(\theta)}{\partial \theta_j^2} \]
satisfies
\[ \nabla^2_{ij}(\theta^0) > 0. \]
For this case, we seek to compute the radius of concavity
\[ r^*_{ij} = \sup\{r : A^{-1}(\theta) \text{ exists and } \nabla^2_{ij}(\theta) > 0 \text{ for all } \theta \in \Theta_r\}. \]
Hence, for uncertainty radii \( r < r^*_{ij} \), \( x_i(\theta) \) is convex with respect to component \( \theta_j \). Similarly, for the case when \( \nabla^2_{ij}(\theta^0) < 0 \), we seek to compute the radius of concavity \( r^*_{ij} \) using the requirement \( \nabla^2_{ij}(\theta^0) < 0 \) for all \( \theta \in \Theta_r \). In this case, for \( r < r^*_{ij} \), \( x_i(\theta) \) is concave with respect to component \( \theta_j \). For the case \( \nabla^2_{ij}(\theta^0) = 0 \), no radius of convexity or concavity is defined; the reader is directed to Section 4 for a discussion of the ramifications of this situation.

2. MAIN RESULT

In this section, the main result of this paper is provided. Namely, an extreme point criterion for determination of the radii of convexity and concavity is given. In the sequel, let \( K_r \) denote the index set for the extreme points (vertices) of hypercube \( \Theta_r \). That is, for \( k \in K_r \), we associate the extreme point \( \theta^k \) of \( \Theta_r \). The result below, stated for radius of convexity, is trivially modified for the radius of concavity.

2.1 Theorem

Consider the set of linear equations
\[ A(\theta)x(\theta) = b(\theta) \]
with solution variable \( x_i(\theta) \) and parameters of interest \( \theta_i \) entering \( A(\theta) \) in a rank-one manner. Then, if \( \nabla^2_{ij}(\theta^0) > 0 \), the radius of convexity is given by
\[ r^*_{ij} = \sup\{r : A^{-1}(\theta^k) \text{ exists and } \nabla^2_{ij}(\theta^k) > 0 \text{ for all } k \in K_r\}. \]

3. PROOF OF THEOREM 2.1

As a first step in the proof, it is noted that the rank-one dependence on \( \theta \) implies that \( \det[A(\theta)] \) is a multilinear function of \( \theta \). Therefore, in view of the well known results on multilinear functions, for example, see (Zadeh and Desoer, 1963), the extreme points of \( \Theta_r \) determine the radius of non-singularity
\[ r_{NS} = \sup\{r : A^{-1}(\theta^k) \text{ exists for all } k \in K_r\}. \]
Accordingly, in the proof of Theorem 2.1, without loss of generality, we assume \( r_{NS} = \infty \).

When \( \theta \) consists of a single parameter \( \theta = \theta_1 \), it is straightforward to show that the solution component \( x_i(\theta) \) can be expressed in the form
\[ x_i(\theta) = \frac{\alpha_\theta + \beta}{\gamma_\theta + \delta} \]
where \( \alpha, \beta, \gamma \) and \( \delta \) are fixed constants depending on \( A_0, d \) and \( e \). Now taking the second derivative for convexity purposes, we obtain
\[ \frac{\partial^2 x_i}{\partial \theta^2} = \frac{-2\gamma(\alpha \delta - \gamma \beta)}{(\gamma_\theta + \delta)^3}. \]

Since invertibility of \( A(\theta) \) guarantees that \( (\gamma \theta + \delta)^3 \) is always of one sign for \( r < r_{NS} \), it is concluded that convexity of \( x_i(\theta) \) is determined by the signs of quantities \( \gamma, (\gamma \theta + \delta) \) and \( \Gamma(x_i) \cong \alpha \delta - \gamma \beta \).

The following lemma, proven in (Ganesan et al., 2001), shows that \( \Gamma(x_i) \) may be expressed as the product of two determinants and a positive constant. We allow the \( \cong \) symbol to denote equality except for a positive multiplicative constant, e.g.,
\[ c^2 \Gamma(x_i) \cong \Gamma(x_1). \]

3.1 Lemma

For the parameterized \( n \)-dimensional linear system of equations
\[ (A_0 + \theta \mathbf{d}_i^T)x = b, \]
with single rank-one uncertainty \( \theta = \theta_1 \),
\[ \Gamma(x_i) \cong \det[A_0 + \theta \mathbf{d}_i] \det \left[ \begin{array}{c} A_0 \mathbf{b} \mathbf{e}^T \end{array} \right] \]
Here, \( [S \leftarrow_i \mathbf{s}] \) denotes the matrix which results when the \( i \)-th column of the matrix \( S \) is replaced with the vector \( s \).

The rank-one uncertainty structure ensures that the above determinants, as well as \( \det[A(\theta)] \) and \( \det[A(\theta) \leftarrow_i b(\theta)] \), are multilinear functions of \( \theta \). The second partial derivative of \( x_i(\theta) \) with respect to \( \theta_i \) may therefore be expressed as a ratio of products of multilinear functions, as detailed in the following lemma which is proven fully in (Ganesan et al., 2001).

3.2 Lemma

For the parameterized \( n \)-dimensional linear equation \( A(\theta)x = b(\theta) \) with rank-one uncertainty
structure, the second partial derivative of \( x_i(\theta) \) with respect to \( \theta_j \) admits a factorization of the form
\[
\frac{\partial^2 x_i}{\partial \theta_j^2} = -2 \frac{g_1(\theta)g_2(\theta)g_3(\theta)}{\det[\mathbf{A}^*(\theta)]},
\]
with \( g_1(\theta), g_2(\theta), g_3(\theta) \) being multilinear functions.

With \( \nabla^2_{ij}(\theta^0) > 0 \), it now follows that \( \nabla^2_{ij}(\theta) \) remains positive for \( \theta \in \Theta \), if and only if each of the multilinear factors \( g_i(\theta) \) has one sign for all \( \theta \in \Theta \). Now, using the basic fact that a multilinear function \( g_i(\theta) \) on a hypercube is positive if and only if the extreme point evaluations \( g_i(\theta^0) \) are positive, for example, see (Zadeh and Desoer, 1963), the formulae for the radii of convexity \( r_{ij}^* \) are readily apparent.

4. MAXIMIZING DISTRIBUTIONS

As indicated in Section 1.2, convexity/concavity information obtained through the use of Theorem 2.1 may be used to determine the joint distribution \( f^* \in \mathcal{F} \) leading to the distributionally robust expected value \( E^* = \max_{f \in \mathcal{F}} E[ x_i(\theta^j) ] \).

That is, the joint distribution \( f^* \) is determined by selecting marginal distributions \( f^*_{ij} \) for the \( \theta_j \) based on the convexity/concavity of \( x_i(\theta) \) with respect to each \( \theta_j \).

More specifically, if \( \nabla^2_{ij}(\theta^0) > 0 \) for all \( \theta \in \Theta \), then \( x_i(\theta) \) is convex in component \( \theta_j \) and therefore \( f^*_{ij} = u \), where \( u \) denotes the uniform distribution over the support interval for \( \theta_j \). If \( \nabla^2_{ij}(\theta^0) < 0 \) for all \( \theta \in \Theta \), then \( x_i(\theta) \) is concave in component \( \theta_j \) and therefore \( f^*_{ij} = \delta \), where \( \delta \) denotes the Dirac delta distribution centered at \( \theta_j^0 \).

When \( \nabla^2_{ij}(\theta^0) = 0 \), Theorem 2.1 cannot be applied to determine convexity or concavity, and \( r_{ij}^* \) is not defined. However, for the special case when \( \nabla^2_{ij}(\theta^0) \equiv 0 \), it is easy to show that the value of \( E[ x_i(\theta^j) ] \) is the same for all admissible \( f^*_{ij} \). This situation occurs, for example, if \( \theta_j \) appears in \( b(\theta) \) and is then absent from \( A(\theta) \) by assumption.

In order to assign the maximizing distributions as above and compute the corresponding distributionally robust maximum expected value \( E^* \), the following two requirements must be met: First, for all \( j \), either \( \nabla^2_{ij}(\theta^0) \neq 0 \), or \( \nabla^2_{ij}(\theta^0) \equiv 0 \). Second, the uncertainty radius \( r \) must not exceed the individual radii of convexity \( r_{ij}^* \) defined above. Now, to obtain the maximum uncertainty radius \( r^* \) for which \( f^* \) is determinable by the method above, let
\[
J \doteq \{ j : \nabla^2_{ij}(\theta^0) \neq 0 \}
\]
and
\[
r^* \doteq \min_{j \in J} r_{ij}^*.
\]

Thus, for \( r < r^* \), distributional robustness can be studied using either the uniform or Dirac delta function distributions as appropriate for each uncertain parameter.

5. ILLUSTRATIVE EXAMPLES

The first example is of low dimension and included for pedagogical purposes, to illustrate the mechanics associated with Theorem 2.1. The second example is also illustrative of the efficacy of Theorem 2.1 in that the dimension of \( \theta \) is a significant factor which complicates the analysis.

5.1 Example (Chain Saw)

This example, adapted from (Bedford and Fowler, 1995) involves a chain saw which is in static equilibrium; see Figure 2. We seek to determine the forces that the operator must apply to cut the log; that is, we seek \( F_x, F_y \), and \( R \). These forces are the solution to the governing equation \( A(\theta)x = b(\theta) \), where

![Fig. 2. Chain saw in static equilibrium](image)

\[
A(\theta) = \begin{bmatrix}
1 & 0 & -1 & 0 & -\frac{1}{2} \\
0 & 1 & -\frac{\sqrt{3}}{2} & 0 & -\frac{\sqrt{3}}{2} \\
0 & 0 & 2(8 + \theta_3) & -\frac{1}{2} (7 + \theta_5)
\end{bmatrix}
\]

\[
b_1(\theta) = 5 + 0.5\theta_2; \\
b_2(\theta) = 10 + \theta_2 - \theta_3; \\
b_3(\theta) = (20 + \theta_4)(6 + 0.5\theta_3) - (21 + \theta_6)(10 + \theta_3) - (1.5 + 0.5\theta_1)(5 + 0.5\theta_2); \\
x_1(\theta) = F_x; \\
x_2(\theta) = F_y; \\
x_3(\theta) = R.
\]

We treat \( \theta \) as a random 8-tuple with probability distribution belonging to the class \( \mathcal{F} \) described in Section 1.2, with \( \theta^0 = 0 \).

We now focus this analysis on \( x_3 \); this is the force with which the operator must “push” to cut the log. Theorem 2.1 is used to determine the...
maximum expected value $E^* = \max_{f \in \mathcal{F}} E[x_3(\theta^f)]$ and also seek various radii of convexity for $x_3$ and $r^*$. Namely, per Section 1.4, we first determine $r_{NS}$. Indeed, with

$$\det \mathbf{A}(\theta) = -\frac{\sqrt{3}}{2} (8 + \theta_5) - \frac{1}{2} (7 + \theta_6)$$

obtained by inspection, the smallest $r$ for which this determinant vanishes on $\Theta_r$ is $r_{NS} \approx 7.6340$.

The value of $r^*$ is determined next: The second partial derivatives of $x_3$ with respect to each $\theta_j$ are $\nabla^2_{3i} (\theta) \equiv 0$ for $j = 1, 4, 7, 8$ and

$$\nabla^2_{35}(\theta) = \frac{3}{2} \frac{b_3(\theta)}{[\det \mathbf{A}(\theta)]^3};$$

$$\nabla^2_{36}(\theta) = \frac{1}{2} \frac{b_5(\theta)}{[\det \mathbf{A}(\theta)]^3}.$$

Since $\nabla^2_{35} (\theta^0) > 0$ and $\nabla^2_{36} (\theta^0) > 0$, the radii of convexity $r_{35}^*$ and $r_{36}^*$ are uniformly distributed over $[-r, r]$ and the remaining $\theta_j$ having any admissible distribution. A closed-form expression for the maximum expected value is thus obtained:

$$E = E^*(r) = \frac{195}{\sqrt{3}} g(r)$$

where

$$g(r) = [\sqrt{3}(8 + r) + 7 + r] \ln[\sqrt{3}(8 + r) + 7 + r]$$

$$- [\sqrt{3}(8 - r) + 7 + r] \ln[\sqrt{3}(8 - r) + 7 + r]$$

$$+ [\sqrt{3}(8 - r) + 7 - r] \ln[\sqrt{3}(8 - r) + 7 - r]$$

$$- [\sqrt{3}(8 + r) + 7 - r] \ln[\sqrt{3}(8 + r) + 7 - r].$$

5.2 Example (Howe Truss)

We consider a Howe truss composed of thirteen elastic members, shown in Figure 3. We seek the horizontal displacement of the endpoint $P$ due to the applied loads $F_i$. It is assumed that each truss member is composed of material with unknown Young’s modulus; the nominal value being that of aluminum ($10^7$ lb/in). The horizontal and vertical members have an uncertain length with nominal value 15 ft. The length of the diagonal members is also uncertain with nominal length $7.5\sqrt{2}$ ft. Each member has an uncertain cross-sectional area with nominal value 1 in$^2$.

The resulting uncertain moduli of elasticity are represented as follows: If the $i$-th member is horizontal or vertical, then the modulus of elasticity $c_k(\theta)$ is linearly approximated about its nominal value by

$$c_k(\theta) = 5.56 \times 10^3 + 13.9\theta_i,$$

where $\theta_i$ is an uncertain parameter with magnitude bound $|\theta_i| \leq 1000$. If the $i$-th member is diagonal, we use the approximation

$$c_k(\theta) = 3.93 \times 10^3 + 9.83\theta_i,$$

where $\theta_i$ is an uncertain parameter with magnitude bound $|\theta_i| \leq 1000$. Thus, each $c_k(\theta)$ may deviate by $\pm 25\%$ of its nominal value. The uncertain applied forces $F_i$ are

$$F_1 = -500 + 0.5\theta_{14}; \quad F_2 = -5000 + \theta_{15};$$

$$F_3 = -500 + 0.5\theta_{16}; \quad F_4 = -500 + \theta_{17};$$

$$F_5 = -500 + 0.5\theta_{18}; \quad F_6 = -1500 + \theta_{19};$$

where $\theta_{14}, \theta_{15}, \theta_{16}, \theta_{17}, \theta_{18}$ and $\theta_{19}$ are uncertain parameters with bound $|\theta_i| \leq 1000$. Finally, the standing assumption is that $\theta$ has probability distribution $f \in \mathcal{F}$ with $\theta^0 = 0$.

To determine the joint distribution which maximizes the expected value of the displacement of point $P$, the displacements of all the joints in Figure 3 are first obtained as the solution of the associated $13 \times 13$ mechanical equation $\mathbf{A}(\theta) \mathbf{x} = \mathbf{b}(\theta)$ with $\theta$ entering in a rank-one manner.

From a practical standpoint, the physical requirement of strict positivity for the moduli of elasticity $c_k(\theta)$ ensures nonsingularity of the system matrix $\mathbf{A}(\theta)$ since

$$\det \mathbf{A} = \prod_{k=1}^{13} c_k(\theta).$$

The linear approximation of $\theta$ leads to $r_{NS} = 4000$. However, since it is now shown that each $r_{ij}^* \ll r_{NS}$, the applicability of Theorem 2.1 is unaffected by the value of $r_{NS}$.

To determine the radii of convexity $r_{ij}^*$ and joint distribution $f^* \in \mathcal{F}$ leading to $E^*$, for each $j$, we compute $\nabla^2_{ij}(\theta)$. For example, we find

![Fig. 3. Truss used in example](image-url)
\[
\n\nabla^2_{11}(\theta) = \frac{2\theta_{12} + 4\theta_{15} + 2\theta_{14} - 25000 + \theta_{16}}{9(3.93 \times 10^4 + 9.83\theta_1)^3}
\]

with the other \(\nabla^2_{ij}(\theta)\) having similar forms. For \(j = 7, 8, 14–19\), it is found that \(\nabla^2_{ij}(\theta) \equiv 0\). The radii of convexity/concavity \(r_{ij}\) are now determined by the extreme point method suggested by Theorem 2.1. The sign of the numerator and denominator of each of the partial derivatives are computed at the extreme points of a hypercube, to determine largest radius for which sign invariance is maintained over the set of extreme points. The radii of convexity are

\[
\begin{align*}
r^*_{12} &\approx 2000; \\
r^*_{15} &\approx 1125; \\
r^*_{16} &\approx 1250;
\end{align*}
\]

\[
\begin{align*}
r^*_{19} &\approx 1250; \\
r^*_{10,10} &\approx 1250; \\
r^*_{11,11} &\approx 1333; \\
r^*_{12} &\approx 1125; \\
r^*_{13,13} &\approx 1125.
\end{align*}
\]

The radii of concavity are

\[
\begin{align*}
r^*_{11} &\approx 2000; \\
r^*_{13} &\approx 2100; \\
r^*_{14} &\approx 2000.
\end{align*}
\]

Indeed, for uncertainty radius \(r = 1000\), the maximizing distribution \(f^*\) is now determined. For the indices \(j = 2, 5, 6, 9, 10–13\), \(f_j^*\) is the uniform distribution, and for \(j = 1, 3, 4\), \(f_j^*\) is the Dirac delta function. For \(j = 7, 8, 14–19\), \(E^*\) is independent of the choice of \(f_j^*\) from the admissible class of marginal distributions. In the analysis to follow, these \(f_j^*\) are chosen to be Dirac delta distributions to simplify computation.

With the maximizing distribution \(f^*\) now defined, \(E^*\) is computed using Monte Carlo sampling with 100,000 sample points. The expected value is found to be \(E^* \approx 0.269\). Convergence of the Monte Carlo estimate with respect to sample size is illustrated in Figure 4; convergence to three significant digits is seen after 10,000 samples. For comparison, the value of \(E[x_1(\theta^0)]\), where \(\theta^0\) denotes the \(m\)-tuple with each component having the uniform distribution, is also computed to be \(E[x_1(\theta^0)] \approx 0.247\). Thus, use of the uniform distribution for every parameter underestimate the maximum expected value of \(x_1\) achievable with \(f \in \mathcal{F}\) by approximately 9%.

6. CONCLUDING REMARKS

Central to the proof of Theorem 2.1 is the fact that the second partial derivative can be factored into a product of multilinear functions. Although a “user” of this result need not actually perform this factorization, it is fundamental to the researcher trying to extend the results of this paper.

Fig. 4. Convergence of Monte Carlo estimate

A fundamental open research problem is motivated by the case when the uncertainty bound exceeds the radius of convexity or concavity. In this case, the probability density function for \(\theta\) leading to the distributionally robust expected value is no longer available. It is felt that a solution for problems of this sort would be important.

7. REFERENCES


