Optimization
Vol. 00, No. 00, Month 200x, 1–13

RESEARCH ARTICLE

A new method to build confidence regions for solutions of stochastic variational inequalities

Shu Lu

(*Received 00 Month 200x; in final form 00 Month 200x*)

Stochastic variational inequalities model a large class of equilibrium problems subject to data uncertainty. The true solution to such a problem is usually estimated by a solution to its sample average approximation (SAA) problem. This paper proposed a new method to build asymptotically exact confidence regions for the true solution that are computable from the SAA solution.

Keywords: variational inequality; stochastic variational inequality; confidence regions; statistical inference; sample average approximation

AMS Subject Classification: 90C33; 90C15

1. Introduction

Variational inequality problems provide a unified framework for modeling equilibrium problems that arise from diverse areas. When the model data are subject to uncertainty, the problem may be formulated as a stochastic variational inequality (SVI). Stochastic variational inequalities are also closely related to stochastic optimization problems.

To define a SVI problem, let $(\Omega, \mathcal{F}, P)$ be a probability space, and ξ be a random vector that is defined on Ω and supported on a closed subset Ξ of $\mathbb{R}^d$. Let O be an open subset of $\mathbb{R}^n$, and F be a measurable function from $O \times \Xi$ to $\mathbb{R}^n$, such that for each $x \in O$ the expectation $f_0(x) = E\|F(x, \xi)\| < \infty$. Let S be a polyhedral convex set in $\mathbb{R}^n$. The SVI problem is to find a point $x \in S \cap O$, called the *true solution*, such that

$$0 \in f_0(x) + N_S(x), \tag{1}$$

where $N_S(x) \subset \mathbb{R}^n$ denotes the normal cone to S at x:

$$N_S(x) = \{v \in \mathbb{R}^n \mid \langle v, s - x\rangle \leq 0 \text{ for each } s \in S\}.$$

We use $\langle \cdot, \cdot \rangle$ to denote the scalar product of two vectors of the same dimension.

In most problems of interest, the function f_0 does not have a closed form expression and it is impossible to evaluate its exact values. A basic approach to solve SVI problems is the *sample average approximation* (SAA) method. Let $\xi^1, \cdots, \xi^N$ be independent and identically distributed (i.i.d.) random variables with distribution

Department of Statistics and Operations Research, University of North Carolina at Chapel Hill

ISSN: 0233-1934 print/ISSN 1029-4945 online

DOI: 10.1080/0233193YYxxxxxxxx
http://www.informaworld.com

same as that of ξ. Define the sample average function $f_N : O \times \Omega \to \mathbb{R}^n$ by

$$f_N(x,\omega) = N^{-1} \sum_{i=1}^{N} F(x, \xi^i(\omega)). \tag{2}$$

The SAA problem is to find a point $x \in S \cap O$, called the *SAA solution*, such that

$$0 \in f_N(x,\omega) + N_S(x). \tag{3}$$

It is well known that the SAA solution will almost surely converge to a true solution as the sample size N goes to ∞, when certain regularity conditions hold [2, 3, 9]. The asymptotic distribution of the SAA solution was obtained in [3, 9]. It was shown in [11] that the SAA solutions converge to the set of true solutions in probability at an exponential rate.

Using the asymptotic distribution of SAA solutions, one can readily obtain an expression for confidence regions of the true solution. However, this expression contains a function that depends discontinuously on the true solution. Due to such discontinuity, the standard procedure of using the SAA solution to replace the true solution in this expression becomes problematic. To overcome this problem, a method was proposed and justified in [4] to build asymptotically exact confidence regions for the true solution, by using a sequence of functions of the SAA solutions to approximate the above discontinuous function evaluated at the true solution.

The objective of this paper is to propose a new method to build asymptotically exact confidence regions for the true solution. This new method is more convenient to implement than the method in [4]. The latter method requires computation of a weighted sum of a family of functions, while the new method only computes a single function in this family. Numerical results show that confidence regions generated by the new method are also more similar to confidence regions that could be generated in "ideal" situations.

The idea of this new method has some similarity to a method proposed in [1, p. 48], with major differences. To obtain a nice approximation of the aforementioned discontinuous function evaluated at the true solution, we search within a certain distance around the SAA solution, to find a point that lies in the lowest dimensional cell (see Section 2 for the definition of cells), and use the function value evaluated at this point as the desired approximation. The method in [1] only considers cells of dimension at least $n-1$, and uses a prespecified distance around the SAA solution. Our method considers all cells and uses a distance that depends on the sample size in a specific manner. The fact that this distance shrinks as the sample size increases enables us to establish a precise limit theorem to justify the method.

Methods developed in [4] and this paper both require samples to be i.i.d. It is possible to extend these methods to non i.i.d samples, by imposing conditions to guarantee certain types of central limit theorems and uniform exponential convergence results to hold.

The paper is organized as follows. Section 2 introduces preliminary results and assumptions under which the proposed method will work. Section 3 presents the method and justifies it. Section 4 specializes the method to complementarity problems and compare numerical results with those in [4].

Throughout this paper, we use $\mathbb{R}^n$, $\mathbb{R}^n_+$ and $\mathbb{R}^n_{++}$ to represent the n-dimensional Euclidean space, its nonnegative orthant, and the interior of the nonnegative orthant. We use $\operatorname{ri} C$ to denote the relative interior of a convex set C. For a convex and closed set $C \subset \mathbb{R}^n$ and a point $z \in \mathbb{R}^n$, $\Pi_C(z)$ denotes the Euclidean projection of z onto C, namely the point in C nearest to z in Euclidean norm. We use

$\|\cdot\|$ to denote the norm of an element in a normed space; unless explicitly stated otherwise, it can be any norm, as long as the same norm is used in all related contexts. We use $\mathcal{N}(\mu, \Sigma)$ to denote a normal random variable with mean μ and covariance matrix Σ. Weak convergence of random variables Y_n to Y will be denoted as $Y_n \Rightarrow Y$. A function $g : \mathbb{R}^n \to \mathbb{R}^m$ is said to be B-differentiable at a point $x_0 \in \mathbb{R}^n$ if there is a positively homogeneous function $dg(x_0) : \mathbb{R}^n \to \mathbb{R}^m$, such that

$$g(x_0 + v) = g(x_0) + dg(x_0)v + o(v).$$

The above function $dg(x_0)$ is the B-derivative of g at x_0.

2. Preliminaries and assumptions

This section is a summary of background information, including preliminary results and the assumptions we make. We refer the reader to [4] for more details.

The variational inequality (1) can be equivalently formulated as an equation. Let $\Pi_S^{-1}(O)$ be the set of points $z \in \mathbb{R}^n$ whose Euclidean projection $\Pi_S(z)$ onto S belongs to O, and define a function $(f_0)_S$ from $\Pi_S^{-1}(O)$ to $\mathbb{R}^n$ as

$$(f_0)_S(z) = f_0(\Pi_S(z)) + (z - \Pi_S(z)). \tag{4}$$

The function $(f_0)_S$ defined above is called the *normal map* induced by f_0 and S. From the above definition, it is not hard to show that a point z satisfies

$$(f_0)_S(z) = 0 \tag{5}$$

if and only if $\Pi_S(z)$ is a solution to (1). Equation (5) is called the *normal map formulation* of (1).

Because S is a polyhedral convex set, the Euclidean projector Π_S is a piecewise affine function on $\mathbb{R}^n$, that coincides with an affine function on each of a family of finitely many n-dimensional polyhedral convex sets. This family of sets is called the *normal manifold* of S, and each set in this family is called an n-cell (the symbol n refers to the dimension of these sets). Each k-dimensional face of each n-cell is called a k-cell in the normal manifold. Each k-cell for $k = 0, \cdots, n$ is called a cell. The relative interiors of all cells in the normal manifold form a partition of $\mathbb{R}^n$. As an example, if $n = 2$ and $S = \mathbb{R}^2_+$, then the normal manifold of S contains nine cells, including four 2-cells (the quadrants), four 1-cells (the half-axes) and one 0-cell (the singleton containing only the origin). More information about normal manifolds can be found in [4–6, 8].

Being piecewise affine, the function Π_S is B-differentiable on $\mathbb{R}^n$. At each $z \in \mathbb{R}^n$, the B-derivative $d\Pi_S(z)$ is a piecewise linear function from $\mathbb{R}^n$ to $\mathbb{R}^n$. Moreover, for all points z in the relative interior of a cell, $d\Pi_S(z)$ is the same function. When z moves from the relative interior of one cell to that of another, $d\Pi_S(z)$ changes abruptly. Thus, $d\Pi_S(\cdot)$ is not continuous on boundaries of the cells in the normal manifold. Let us use $S = \mathbb{R}^2_+$ again to illustrate. At any point $z \in \operatorname{int} S$, the B-derivative $d\Pi_S(z)$ is the identity map on $\mathbb{R}^2$. At $z \in \{0\} \times \mathbb{R}_{++}$, $d\Pi_S(z)$ is a piecewise linear function with two pieces. The B-derivative at the origin, $d\Pi_S(0)$, is a piecewise linear function with four pieces. Section 4 contains a detailed discussion about $S = \mathbb{R}^n_+$, where n is any integer.

The assumptions needed for the method of this paper to work are the same as those in [4]; we present them below for the sake of completeness. In the rest of this

paper, we will use X to denote a nonempty compact subset of O. Let $C^1(X, \mathbb{R}^n)$ be the Banach space of continuously differentiable mappings $f : X \to \mathbb{R}^n$, equipped with the norm

$$\|f\|_{1,X} = \sup_{x \in X} \|f(x)\| + \sup_{x \in X} \|df(x)\|. \tag{6}$$

For a point $x \in S$, we use $T_S(x)$ to denote the tangent cone to S at x. Since S is a polyhedral convex set, the following definition applies:

$$T_S(x) = \{v \in \mathbb{R}^n \mid \text{ there exists } t \in \mathbb{R}_{++} \text{ such that } x + tv \in S\}.$$

Assumption 2.1 below will guarantee that the function f_0 belongs to $C^1(X, \mathbb{R}^n)$ and that f_N converges to f_0 almost surely as an element of $C^1(X, \mathbb{R}^n)$. Assumption 2.2 ensures that (1) has a locally unique solution under sufficiently small perturbation of f_0. The set K defined there is called the critical cone to S associated with z_0. Assumption 2.3 is a standard nondegeneracy condition. Assumption 2.4 implies that f_N converges to f_0 in probability at an exponential rate; that is, there exist positive real numbers β_1, μ_1, M_1 and σ_1, such that the following holds for each $\epsilon > 0$ and each N:

$$\text{Prob}\left\{\|f_N - f_0\|_{1,X} \geq \epsilon\right\} \leq \beta_1 \exp\{-N\mu_1\} + \frac{M_1}{\epsilon^n} \exp\left\{-\frac{N\epsilon^2}{\sigma_1}\right\}. \tag{7}$$

Assumption 2.1 (a) $E\|F(x, \xi)\|^2 < \infty$ for all $x \in O$.
(b) The map $x \mapsto F(x, \xi(\omega))$ is continuously differentiable on O for a.e. $\omega \in \Omega$, and $E\|dF_x(x, \xi)\|^2 < \infty$ for all $x \in O$.
(c) There exists a square integrable random variable C such that

$$\|F(x, \xi(\omega)) - F(x', \xi(\omega))\| + \|dF(x, \xi(\omega)) - dF(x', \xi(\omega))\| \leq C(\omega)\|x - x'\|,$$

for all $x, x' \in O$ and a.e. $\omega \in \Omega$.

Assumption 2.2 Suppose that x_0 solves the variational inequality (1) and that x_0 belongs to the interior of X. Let $z_0 = x_0 - f_0(x_0)$, $L = df_0(x_0)$, $K = T_S(x_0) \cap \{z_0 - x_0\}^\perp$, and assume that the normal map L_K induced by L and K is a homeomorphism from $\mathbb{R}^n$ to $\mathbb{R}^n$.

Assumption 2.3 Let Σ_0 denote the covariance matrix of $F(x_0, \xi)$. Suppose that the determinant of Σ_0 is strictly positive.

Assumption 2.4 (a) For each $t \in \mathbb{R}^n$ and $x \in X$, let

$$M_x(t) = E\big[\exp\{\langle t, F(x, \xi) - f_0(x)\rangle\}\big]$$

be the moment generating function of the random variable $F(x, \xi) - f_0(x)$. Assume

(1) There exists $\zeta > 0$ such that $M_x(t) \leq \exp\{\zeta^2\|t\|^2/2\}$ for every $x \in X$ and every $t \in \mathbb{R}^n$.
(2) There exists a nonnegative random variable κ such that

$$\|F(x, \xi(\omega)) - F(x', \xi(\omega))\| \leq \kappa(\omega)\|x - x'\| \tag{8}$$

for all $x, x' \in O$ and almost every $\omega \in \Omega$.

(3) The moment generating function of κ is finite valued in a neighborhood of zero.

(b) For each $T \in \mathbb{R}^{n\times n}$ and $x \in X$, let

$$\mathcal{M}_x(T) = E\big[\exp\{\langle T, d_xF(x,\xi) - df_0(x)\rangle\}\big]$$

be the moment generating function of the random variable $d_xF(x,\xi) - df_0(x)$. Assume

(1) There exists $\varsigma > 0$ such that $\mathcal{M}_x(T) \le \exp\{\varsigma^2\|T\|^2/2\}$ for every $x \in X$ and every $T \in \mathbb{R}^{n\times n}$.
(2) There exists a nonnegative random variable ν such that

$$\|d_xF(x,\xi(\omega)) - d_xF(x',\xi(\omega))\| \le \nu(\omega)\|x - x'\|$$

for all $x, x' \in O$ and almost every $\omega \in \Omega$.
(3) The moment generating function of ν is finite valued in a neighborhood of zero.

Assumption 2.4 above is a special case of a more general condition, which implies

$$\text{Prob}\,\{\|f_N - f_0\|_{1,X} \ge \epsilon\} \le C(\epsilon)\exp\{-N\beta(\epsilon)\}$$

for some functions $C(\epsilon)$ and $\beta(\epsilon)$, see [10–12]. To prove the main result of this paper we need $C(\epsilon)$ and $\beta(\epsilon)$ to be in explicit forms, so we require Assumption 2.4 to hold. Assumption 2.4 is satisfied, for example, if $F(x,\xi)$ is a bounded random variable uniformly in x, or if $F(x,\xi)$ is a symmetric random variable with its norm being stochastically dominated by the absolute value of some normal random variable [4].

The following theorem is adapted from [4, Theorem 5.1].

Theorem 2.5 : *Suppose that Assumptions 2.1 and 2.2 hold. Let Y_0 be a normal random vector in $\mathbb{R}^n$ with zero mean and covariance matrix Σ_0. Then there exist neighborhoods X_0 of x_0 and Z of z_0 such that the following hold. For almost every $\omega \in \Omega$, there exists an integer N_ω, such that for each $N \ge N_\omega$, the equation*

$$(f_N)_S(z) = 0 \tag{9}$$

has a unique solution z_N in Z, and the variational inequality (3) has a unique solution in X_0 given by $x_N = \Pi_S(z_N)$. Moreover, $\lim_{N\to\infty} z_N = z_0$ almost surely,

$$\sqrt{N}(z_N - z_0) \Rightarrow (L_K)^{-1}(Y_0), \tag{10}$$

and

$$\sqrt{N}L_K(z_N - z_0) \Rightarrow Y_0. \tag{11}$$

Suppose in addition that Assumption 2.4 holds. Then there exist positive real numbers ϵ_0, β_0, μ_0, M_0 and σ_0, such that the following holds for each $\epsilon \in (0, \epsilon_0]$

and each N:

$$\begin{aligned}&\text{Prob}\left\{\|x_N - x_0\| < \epsilon\right\} \geq \text{Prob}\left\{\|z_N - z_0\| < \epsilon\right\}\\ \geq &1 - \beta_0 \exp\{-N\mu_0\} - \frac{M_0}{\epsilon^n} \exp\left\{-\frac{N\epsilon^2}{\sigma_0}\right\}.\end{aligned} \tag{12}$$

3. A new method to build confidence regions

Using the asymptotic distribution of z_N as given in (11), we can obtain an expression for confidence regions of z_0. To this end, suppose Assumption 2.3 holds in addition to Assumptions 2.1 and 2.2. Then (11) implies

$$\sqrt{N}\Sigma_0^{-1/2} L_K(z_N - z_0) \Rightarrow \mathcal{N}(0, I_n). \tag{13}$$

Consequently, when N is large, the 2-norm of the left hand side of (13) approximately follows the χ^2 distribution with n degrees of freedom, and the set

$$\{z \in \mathbb{R}^n \mid N\big[L_K(z_N - z)\big]^T \Sigma_0^{-1} \big[L_K(z_N - z)\big] \leq \chi_n^2(\alpha)\} \tag{14}$$

defines an approximate $(1-\alpha)100\%$ confidence region for z_0, where $\chi_n^2(\alpha)$ is defined to be the number that satisfies $P(U > \chi_n^2(\alpha)) = \alpha$ for a χ^2 random variable U with n degrees of freedom.

By their definitions, Σ_0 and L_K depend on the true solutions x_0 and z_0, which are unknown. We will use the sample covariance matrix of $\{F(x_N, \xi^i)\}_{i=1}^N$, denoted by Σ_N in the rest of this paper, to approximate Σ_0. As for L_K, the situation is more complicated; we will devote the rest of this section to construction of an appropriate approximation of it.

Under Assumption 2.1, f_0 is continuously differentiable on O (see [4, Theorem 4.1]). As noted in Section 2, Π_S is B-differentiable on $\mathbb{R}^n$. By the chain rule of B-differentiability, the normal map $(f_0)_S$ is B-differentiable on $\Pi_S^{-1}(O)$, with

$$d(f_0)_S(z)(h) = df_0(\Pi_S(z))(d\Pi_S(z)(h)) + h - d\Pi_S(z)(h) \tag{15}$$

for each $z \in \Pi_S^{-1}(O)$ and each $h \in \mathbb{R}^n$. The normal map L_K that appears in (14) is known to be exactly $d(f_0)_S(z_0)$ [7]. However, as noted in Section 2, $d\Pi_S(z)$ is not continuous with respect to z on boundaries of the cells in the normal manifold of S, so $d(f_0)_S(z)$ is not continuous with respect to z on the boundaries either. It is therefore problematic to use $d(f_0)_S(z_N)$ to approximate $d(f_0)_S(z_0)$.

To build a nice approximation of $d(f_0)_S(z_0)$, we start by building a nice approximation of $d\Pi_S(z_0)$. Denote the cells in the normal manifold of S by $C_1, \cdots, C_l$, where l is the number of cells. Recall from Section 2 that the relative interiors of these cells form a partition of $\mathbb{R}^n$. More specifically, for each $z \in \mathbb{R}^n$, exactly one cell contains z in its relative interior, and this cell has the smallest dimension among all cells containing z [4, Propostion 5.1]. For each cell C_i, define a function $d_i : \mathbb{R}^n \to \mathbb{R}$ by

$$d_i(z) = d(z, C_i) = \min_{x \in C_i} \|z - x\|, \tag{16}$$

and a function $\Psi_i : \mathbb{R}^n \to \mathbb{R}^n$ by

$$\Psi_i(\cdot) = d\Pi_S(z)(\cdot) \text{ for any } z \in \operatorname{ri} C_i. \tag{17}$$

Note that the norm that appears in (16) can be any norm for vectors in $\mathbb{R}^n$, and that the point z in (17) can be any point in the relative interior of C_i.

Next, define a function g from the set of integers to $\mathbb{R}$, such that

(i) $g(N) > 0$ for each $N \in \mathbb{N}$.
(ii) $\lim_{N\to\infty} g(N) = \infty$.
(iii) $\lim_{N\to\infty} \frac{N}{g(N)^2} = \infty$.
(iv) $\lim_{N\to\infty} g(N)^n \exp\{-\theta_0 \frac{N}{(g(N))^2}\} = 0$ for $\theta_0 = \min\{\frac{1}{4\sigma_0}, \frac{1}{4\sigma_1}, \frac{1}{4\sigma_0(E(C))^2}\}$, where σ_1, σ_0 are as in (7) and (12) respectively and C is as in Assumption 2.1.
(v) $\lim_{N\to\infty} \frac{N^{n/2}}{g(N)^n} \exp\left\{-\theta g(N)^2\right\} = 0$ for each positive real number θ.

Note that $g(N) = N^p$ for any $p \in (0, 1/2)$ satisfies (i) – (v) above. Positive linear combinations of such functions also satisfy all the above conditions. Another example is $g(N) = \frac{N^{1/2}}{\sqrt{m \log N}}$, where $m > \frac{n}{2\theta_0}$.

Now, for each integer N and any point $z \in \mathbb{R}^n$, choose an index i_0 by letting C_{i_0} be a cell that has the smallest dimension among all cells C_i such that $d_i(z) \leq 1/g(N)$, and define a function $\Lambda_N(z) : \mathbb{R}^n \to \mathbb{R}^n$ by

$$\Lambda_N(z)(h) = \Psi_{i_0}(h) \text{ for each } h \in \mathbb{R}^n. \tag{18}$$

The theorem below shows that $\Lambda_N(z_N)$ provides a nice approximation for $d\Pi_S(z_0)$.

Theorem 3.1 : *Suppose that Assumptions 2.1, 2.2 and 2.4 hold. For each $N \in \mathbb{N}$, let $\Lambda_N(z_N)$ be as defined in* (18) *with z_N in place of z. Then*

$$\lim_{N\to\infty} \text{Prob}\left[\Lambda_N(z_N)(h) = d\Pi_S(z_0)(h) \text{ for all } h \in \mathbb{R}^n\right] = 1. \tag{19}$$

Proof: Let $C_{i(1)}, \cdots, C_{i(q)}$ be the cells in the normal manifold of S that contain z_0, and let $C_{i(q)}$ be the one that contains z_0 in its relative interior. By the remark right above (16), the dimension of $C_{i(q)}$ is strictly smaller than $C_{i(1)}, \cdots, C_{i(q-1)}$. The definition of Ψ_i in (17) implies that

$$d(\Pi_S)(z_0)(h) = \Psi_{i(q)}(h) \text{ for each } h \in \mathbb{R}^n. \tag{20}$$

By Theorem 2.5, there exist positive constants ϵ_0, β_0, μ_0, M_0 and σ_0, such that (12) holds for each $\epsilon \in (0, \epsilon_0]$ and each N. Let γ_0 be the minimum of $d_i(z_0)$ among all cells C_i not containing z_0; we have $\gamma_0 > 0$. Define $\gamma = \min(\gamma_0, 2\epsilon_0)$. The definition of g ensures that $\lim_{N\to\infty} g(N) = \infty$, so there exists an integer N_0, such that $g(N) \geq \max(2/\gamma, 1)$ for each $N \geq N_0$.

An application of (12) to $\epsilon = \gamma/2$ gives

$$\text{Prob}\left\{\|z_N - z_0\| < \frac{\gamma}{2}\right\} \geq 1 - \beta_0 \exp\{-N\mu_0\} - \frac{2^n M_0}{\gamma^n} \exp\left\{-\frac{N\gamma^2}{4\sigma_0}\right\}$$

for each N. For each cell C_i not containing z_0, we have

$$d_i(z_N) \geq d_i(z_0) - \|z_N - z_0\| \geq \gamma - \|z_N - z_0\|.$$

Consequently, for each $N \geq N_0$,

$$\begin{aligned} &\text{Prob}\left\{d_i(z_N) > 1/g(N) \text{ for all } C_i \text{ not containing } z_0\right\} \\ \geq &\,\text{Prob}\left\{\|z_N - z_0\| < \gamma/2\right\} \\ \geq &1 - \beta_0 \exp\{-N\mu_0\} - \frac{2^n M_0}{\gamma^n} \exp\left\{-\frac{N\gamma^2}{4\sigma_0}\right\}. \end{aligned} \tag{21}$$

Next, consider a cell that contains z_0, that is, one of the cells $C_{i(1)}, \cdots, C_{i(q)}$. The following inequality holds for each N and each $j = 1, \cdots, q$,

$$0 \leq d_{i(j)}(z_N) = d(z_N, C_{i(j)}) \leq \|z_N - z_0\|,$$

which implies, for each $N \geq N_0$,

$$\begin{aligned} &\text{Prob}\left\{d_{i(j)}(z_N) < 1/(2g(N)) \text{ for all } j = 1, \cdots, q\right\} \\ \geq &\,\text{Prob}\left\{\|z_N - z_0\| < 1/(2g(N))\right\} \\ \geq &1 - \beta_0 \exp\{-N\mu_0\} - 2^n M_0 g(N)^n \exp\left\{-\frac{N}{4\sigma_0 g(N)^2}\right\}, \end{aligned} \tag{22}$$

where the last inequality follows from (12).

Now, according to (21) and (22), for each $N \geq N_0$, the probability that the family of cells C_i that satisfy $d_i(z_N) \leq 1/g(N)$ is exactly $\{C_{i(1)}, \cdots, C_{i(q)}\}$ is at least

$$1 - 2\beta_0 \exp\{-N\mu_0\} - \frac{2^n M_0}{\gamma^n} \exp\left\{-\frac{N\gamma^2}{4\sigma_0}\right\} - 2^n M_0 g(N)^n \exp\left\{-\frac{N}{4\sigma_0 g(N)^2}\right\}. \tag{23}$$

The expression above converges to 1, according to item (iv) in the definition of $g(N)$. If we let C_{i_0} be a cell that has the smallest dimension among all cells C_i that satisfy $d_i(z_N) \leq 1/g(N)$, then the probability for the index i_0 to be exactly the same as $i(q)$ is greater than or equal to the quantity in (23). This completes the proof in view of (20) and the definition of $\Psi_N(z_N)$. □

According to Theorem 3.1, the probability for $\Lambda_N(z_N)$ to be exactly $d\Pi_S(z_0)$ converges to 1 as N goes to infinity. Thus, with large N, $\Lambda_N(z_N)$ is a good approximation for $d\Pi_S(z_0)$. In comparison to [4], $\Lambda_N(z_N)$ defined here is easier to compute and has simpler structure, since it coincides with a single function in the family $\{\Psi_i\}$.

Next, we define for each $N \in \mathbb{N}$ a function $\Phi_N : \Pi_S^{-1}(O) \times \mathbb{R}^n \times \Omega \to \mathbb{R}^n$ by

$$\Phi_N(z, h, \omega) = df_N(\Pi_S(z))(\Lambda_N(z)(h)) + h - \Lambda_N(z)(h). \tag{24}$$

For convenience we will write $\Phi_N(z_N(\omega), h, \omega)$ as $\Phi_N(z_N)(h)$ in the rest of this paper. The corollaries below show that $\Phi_N(z_N)$ provides a good approximation for $d(f_0)_S(z_0)$, and that the random vector $\sqrt{N}\Sigma_N^{-1/2}\Phi_N(z_N)(z_N - z_0)$ approximately follows the standard normal distribution.

Corollary 3.2: *Suppose that Assumptions 2.1, 2.2 and 2.4 hold. For each $N \in \mathbb{N}$,*

let Φ_N be as defined in (24)*. Then, there exists a positive real number ϕ, such that*

$$\lim_{N\to\infty} \text{Prob}\left[\sup_{h\in\mathbb{R}^n} \frac{\|\Phi_N(z_N)(h) - d(f_0)_S(z_0)(h)\|}{\|h\|} < \frac{\phi}{g(N)}\right] = 1. \tag{25}$$

Proof: The proof follows the same arguments in the proof of [4, Corollaries 5.2]. □

Corollary 3.3: *Suppose that Assumptions 2.1, 2.2 and 2.4 hold. Then*

$$\sqrt{N}\Phi_N(z_N)(z_N - z_0) \Rightarrow \mathcal{N}(0, \Sigma_0). \tag{26}$$

If Assumption 2.3 holds additionally, then

$$\sqrt{N}\Sigma_N^{-1/2}\Phi_N(z_N)(z_N - z_0) \Rightarrow \mathcal{N}(0, I_n). \tag{27}$$

Proof: The proof follows the same arguments in the proofs of [4, Corollary 5.2 and Theorem 3.1]. □

As a result of Corollary 3.3, when N is large, the 2-norm of the left hand side of (27) approximately follows the χ^2 distribution with n degrees of freedom, and the set

$$\{z \in \mathbb{R}^n \mid N\big[\Phi_N(z_N)(z_N - z)\big]^T \Sigma_N^{-1}\big[\Phi_N(z_N)(z_N - z)\big] \le \chi_n^2(\alpha)\} \tag{28}$$

defines an approximate $(1-\alpha)100\%$ confidence region for z_0. The expression (28) is different from (14), in that it does not depend on z_0 and is therefore computable.

4. Implementation in complementarity problems

An important subclass of variational inequalities is the family of complementarity problems, in which the set S takes the form of $\mathbb{R}^p_+ \times \mathbb{R}^{n-p}$. This section implements the confidence region technique to stochastic complementarity problems of the form

$$0 \in f_0(x) + N_{\mathbb{R}^p_+ \times \mathbb{R}^{n-p}}(x). \tag{29}$$

The normal manifold of $\mathbb{R}^p_+ \times \mathbb{R}^{n-p}$ consists of a total of 3^p cells. Associated with each partition (I_0, I_+, I_-) of the index set $\{1, \cdots, p\}$ is a cell

$$C(I_0, I_+, I_-) = \{x \in \mathbb{R}^n \mid x_i = 0, i \in I_0;\ x_i \ge 0, i \in I_+;\ x_i \le 0, i \in I_-\}. \tag{30}$$

The dimension of $C(I_0, I_+, I_-)$ is $n - |I_0|$, where $|I_0|$ is the cardinality of I_0. For a point z that belongs to the relative interior of $C(I_0, I_+, I_-)$, the B-derivative of the Euclidean projector $\Pi_{\mathbb{R}^p_+ \times \mathbb{R}^{n-p}}$ is given by

$$(d\Pi_{\mathbb{R}^p_+ \times \mathbb{R}^{n-p}}(z)(h))_i = \begin{cases} h_i \text{ if } i \in I_0 \text{ and } h_i \ge 0, \\ 0 \ \text{ if } i \in I_0 \text{ and } h_i \le 0, \\ h_i \text{ if } i \in I_+, \\ 0 \ \text{ if } i \in I_-, \\ h_i \text{ if } i \in \{p+1, \cdots, n\} \end{cases} \tag{31}$$

for each $h \in \mathbb{R}^n$; see [4, Lemma 6.1]. Proposition 4.1 below provides a formula for computing Λ_N.

Proposition 4.1: *Let $N \in \mathbb{N}$, $z \in \mathbb{R}^n$ and let the cell $C(I_0^*, I_+^*, I_-^*)$ be a cell of the smallest dimension among all cells $C(I_0, I_+, I_-)$ such that $d(z, C(I_0, I_+, I_-)) \leq 1/g(N)$. Then, the function Λ_N defined in (18) with $S = \mathbb{R}^p_+ \times \mathbb{R}^{n-p}$ is a piecewise linear function from $\mathbb{R}^n$ to $\mathbb{R}^n$. On each set of the form $Y \times \mathbb{R}^{n-p}$, with Y being an orthant of $\mathbb{R}^p$, $\Lambda_N(z)(\cdot)$ is represented by the following $n \times n$ diagonal matrix:*

$$\Lambda_N(z, Y) = \begin{bmatrix} \lambda_1 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_p & 0 & \cdots & 0 \\ 0 & \cdots & 0 & 1 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & 0 & \cdots & 1 \end{bmatrix},$$

where for each $i = 1, \cdots, p$,

$$\lambda_i = \begin{cases} 1 & \text{if } i \in I_+^* \text{ or } (i \in I_0^* \text{ and the } i\text{th components of vectors in } Y \text{ are nonnegative}) \\ 0 & \text{if } i \in I_-^* \text{ or } (i \in I_0^* \text{ and the } i\text{th components of vectors in } Y \text{ are nonpositive}). \end{cases}$$

Proof: Let z^* be a point that belongs to the relative interior of the cell $C(I_0^*, I_+^*, I_-^*)$. The way Λ_N is defined in (18) gives $\Lambda_N(z)(\cdot) = d\Pi_{\mathbb{R}^p_+ \times \mathbb{R}^{n-p}}(z^*)(\cdot)$. Applying the formula (31) with I_0^*, I_+^*, I_-^* in place of I_0, I_+, I_-, we find that

$$(\Lambda_N(z)(h))_i = \begin{cases} h_i \text{ if } i \in I_0^* \text{ and } h_i \geq 0, \\ 0 \text{ if } i \in I_0^* \text{ and } h_i \leq 0, \\ h_i \text{ if } i \in I_+^*, \\ 0 \text{ if } i \in I_-^*, \\ h_i \text{ if } i \in \{p+1, \cdots, n\} \end{cases}$$

for each $h \in \mathbb{R}^n$. Putting this in matrix representation proves this proposition. □

In view of Proposition 4.1, under Assumption 2.1 the function $\Phi_N(z)(\cdot)$ defined in (24) is almost surely a piecewise linear function from $\mathbb{R}^n$ to $\mathbb{R}^n$. On each set of the form $Y \times \mathbb{R}^{n-p}$ with Y being an orthant of $\mathbb{R}^p$, it is represented by the following matrix:

$$df_N(\Pi_S(z))\Lambda_N(z, Y) + I_n - \Lambda_N(z, Y),$$

where $\Lambda_N(z, Y)$ is as defined in Proposition 4.1.

We implement this method to the numerical example in [4] to compare the performance of both methods. The example is a stochastic linear complementarity problem

$$0 \in E[F(x, \xi)] + N_{\mathbb{R}^2_+}(x), \tag{32}$$

in which $n = 2$, $d = 6$, and $F : \mathbb{R}^2 \times \mathbb{R}^6 \to \mathbb{R}^2$ is defined by

$$F(x, \xi) = \begin{bmatrix} \xi_1 & \xi_2 \\ \xi_3 & \xi_4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} \xi_5 \\ \xi_6 \end{bmatrix}. \tag{33}$$

The random vector ξ follows the uniform distribution over the box $[0,2]\times[0,1]\times[0,2]\times[0,4]\times[-1,1]\times[-1,1]$. For this example, the true solution x_0 for (32) is computable and is given by $x_0 = 0$. As a result, $z_0 = x_0 - E[F(x_0,\xi)] = 0$.

With sample size $N = 10$, we obtain the sample average function

$$f_{10}(x) = M_{10}x + b_{10} = \begin{bmatrix} 0.9292 & 0.5400 \\ 0.7536 & 2.1111 \end{bmatrix} x + \begin{bmatrix} -0.1319 \\ -0.2906 \end{bmatrix},$$

which leads to $x_{10} = (0.0782, 0.1097)$ and $z_{10} = x_{10} - f_{10}(x_{10}) = (0.0782, 0.1097)$. We choose $g(N) = N^{1/3}$ to find $1/g(10) = 0.4642$. The distance from z_{10} to any of the nine cells in the normal manifold of $\mathbb{R}^2$ is not larger than 0.4642. The cell with the smallest dimension among all these cells is the set $\{0\}$, which corresponds to $I_0^* = \{1,2\}$ and $I_+^* = I_-^* = \emptyset$. Thus, $\Lambda_{10}(z_{10})(\cdot)$ is exactly $d\Pi_{\mathbb{R}^2_+}(0)$, which is a piecewise linear map represented by matrices

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \text{ and } \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

in orthants $\mathbb{R}^2_+$, $\mathbb{R}_+\times\mathbb{R}_-$, $\mathbb{R}_-\times\mathbb{R}_+$ and $\mathbb{R}^2_-$ respectively. Accordingly, $\Phi_{10}(z_{10})(\cdot)$ is a piecewise linear map represented by matrices

$$\begin{bmatrix} 0.9292 & 0.5400 \\ 0.7536 & 2.1111 \end{bmatrix}, \begin{bmatrix} 0.9292 & 0 \\ 0.7536 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0.5400 \\ 0 & 2.1111 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

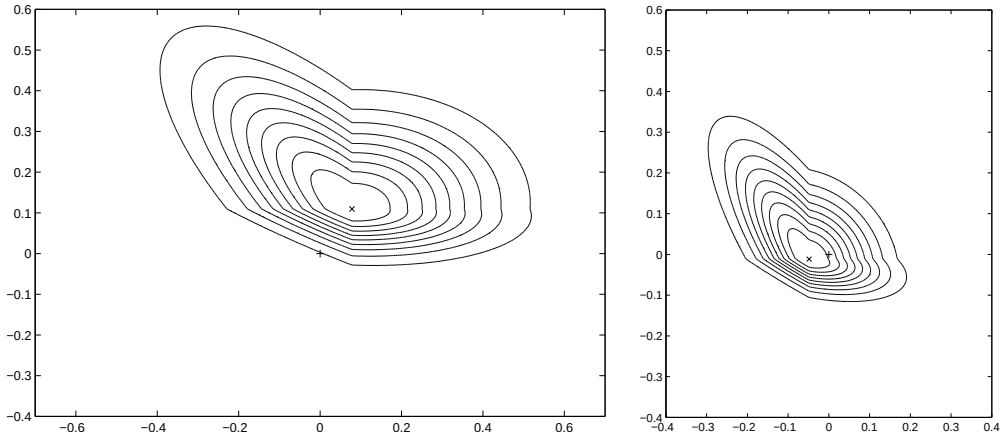
in corresponding orthants. In addition, we obtain the following sample covariance matrix of $F(x_{10},\xi)$:

$$\Sigma_{10} = \begin{bmatrix} 0.4169 & 0.0137 \\ 0.0137 & 0.1865 \end{bmatrix}.$$

Sets in (28) now have explicit expressions. Figure 1(a) shows boundaries of these sets generated using the contour function of Matlab. These regions are centered around z_{10}, marked by '$\times$' in the graph. From the innermost to the outermost, the curves correspond to boundaries of confidence regions for z_0 at levels $0.1,\cdots,0.9$ respectively. The true solution z_0, known to be 0 for this example, is marked by '+' and lies just beyond the 90% confidence region. Figure 1(b) shows confidence regions for z_0 obtained from a different SAA problem with sample size $N = 30$, $x_{30} = 0$ and $z_{30} = (-0.0483, -0.0114)$. The true solution z_0 lies within the 20% region.

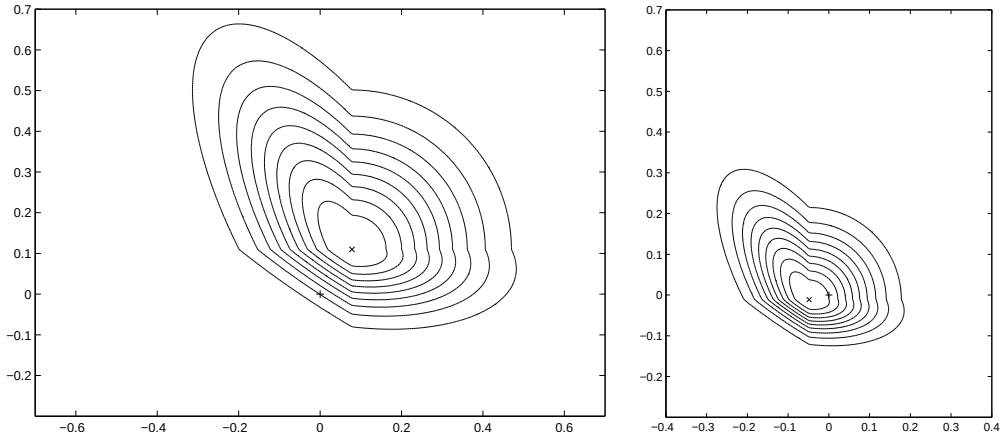
Since the true values of z_0 and Σ_0 are known for this example, the ideal confidence regions for z_0 given in (14) are computable. Figure 2 displays these regions centered around z_{10} and z_{30}. To compare sizes of regions in Figure 1 with those in Figure 2, we find the areas enclosed by the outermost curves in Figures 1(a) and 2(a) to be 0.3454 and 0.3988 respectively, with the difference being about 13% of 0.3988. The corresponding areas in 1(b) and 2(b) are 0.1305 and 0.1329 respectively, which are different by about 2% of 0.1329.

Note that the shapes and sizes of confidence regions generated by the proposed method are more similar to the ideal confidence regions, as comparing to the confidence regions generated in [4]. The reason is that the function $\Lambda_N(z_N)$ here is with high probability exactly equal to $d\Pi_S(z_0)$, whereas $\Lambda_N(z_N)$ defined in [4] is in general different from $d\Pi_S(z_0)$.



(a) $N = 10, z_{10} \approx (0.08, 0.11)$ (b) $N = 30, z_{30} \approx (-0.05, -0.01)$

Figure 1. Confidence regions for z_0 constructed by the proposed method



(a) $N = 10, z_{10} \approx (0.08, 0.11)$ (b) $N = 30, z_{30} \approx (-0.05, -0.01)$

Figure 2. Confidence regions for z_0 constructed from true f_0, z_0 and Σ_0

Acknowledgement

Research of the author was supported by National Science Foundation under grants DMS-0807893 and DMS-1109099. The author thanks Professor Stephen M. Robinson for introducing her to this subject.

References

[1] Demir, M. C. (2000). *Asymptotics and confidence regions for stochastic variational inequalities*. Ph.D. Dissertation, Department of Industrial Engineering, University of Wisconsin–Madison, Madison, WI.

[2] Gürkan, G., Özge, A. Y., and Robinson, S. M. (1999). Sample-path solution of stochastic variational inequalities. *Mathematical Programming*, 84:313–333.

[3] King, A. J. and Rockafellar, R. T. (1993). Asymptotic theory for solutions in statistical estimation and stochastic programming. *Mathematics of Operations Research*, 18:148–162.

[4] Lu, S. and Budhiraja, A. (2011). Confidence regions for stochastic variational ienqualities. Submitted for publication.

[5] Lu, S. and Robinson, S. M. (2008). Variational inequalities over perturbed polyhedral convex sets. *Mathematics of Operations Research*, 33:689–711.

[6] Robinson, S. M. (1992). Normal maps induced by linear transformations. *Mathematics of Operations Research*, 17:691–714.

[7] Robinson, S. M. (1995). Sensitivity analysis of variational inequalities by normal-map techniques. In Giannessi, F. and Maugeri, A., editors, *Variational Inequalities and Network Equilibrium Problems*, pages 257–269, New York. Plenum Press.

[8] Scholtes, S. (1994). *Introduction to Piecewise Differentiable Equations*. Habilitationsschrift, Insti-

tut für Statistik und mathematische Wirtschaftstheorie, Universität Fridericiana Karlsruhe, Karlsruhe, Germany.

[9] Shapiro, A., Dentcheva, D., and Ruszczyński, A. (2009). *Lectures on Stochastic Programming: Modeling and Theory*. Society for Industrial and Applied Mathematics and Mathematical Programming Society, Philadelphia, PA.

[10] Shapiro, A. and Xu, H. (2008). Stochastic mathematical programs with equilibrium constraints, modelling and sample average approximation. *Optimization*, 57:395–418.

[11] Xu, H. (2010a). Sample average approximation methods for a class of stochastic varational inequality problems. *Asia-Pacific Journal of Operational Research*, 27(1):103–119.

[12] Xu, H. (2010b). Uniform exponential convergence of sample average random functions under general sampling with applications in stochastic programming. *Journal of Mathematical Analysis and Applications*, 368:692–710.