Proceedings of the 2006 Winter Simulation Conference
L. F. Perrone, F. P. Wieland, J. Liu, B. G. Lawson, D. M. Nicol, and R. M. Fujimoto, eds.

GRADIENT-BASED SIMULATION OPTIMIZATION

Sujin Kim

School of Industrial Engineering
Purdue University
West Lafayette, IN 47907, U.S.A.

ABSTRACT

We present a review of methods for simulation optimization. In particular, we focus on gradient-based techniques for continuous optimization. We demonstrate the main concepts using as an example the multidimensional newsvendor problem. We also discuss mathematical techniques and results that are useful in verifying and analyzing the simulation optimization procedures.

1 INTRODUCTION

Recent advances in computing power have inspired great interest in simulation optimization. Many techniques for simulation optimization have been developed, but their effectiveness varies greatly depending on the details of the problem under consideration: the underlying structure of the decision variables, objective function, and constraints.

The general optimization problem we consider in this paper has the following form:

$$\min_{\theta\in\Theta} f(\theta) = E[f(\theta,\xi)], \tag{1}$$

for some random variable ξ and parameter $\theta \in \Theta$, where $\Theta \in \mathbb{R}^p$ is a set of possible values of the parameter θ. Note that (1) is a stochastic optimization problem. We assume that the function $f(\theta)$ is differentiable, and can only be evaluated using Monte Carlo simulation. Thus, standard gradient-based stochastic optimization algorithms can be used to compute an (approximate) minimizer of (1).

The purpose of this paper is to provide a tutorial in simulation optimization methods for solving problems like (1). We will consider only those simulation optimization problems with continuous variables and differentiable objective functions. Furthermore, in our discussion of those problems, we will focus only on gradient-based procedures. We will not cover some other important classes of methods for simulation optimization, including response surface methodology, ranking and selection, and random search. For excellent comprehensive reviews of literature on simulation optimization, which cover these topics, see Andradóttir (1998), Azadivar (1999), Fu (1994), Fu (2002), Fu, Glover, and April (2005), Henderson and Nelson (2006).

The emphasis is on the mathematical techniques that can be used to answer the following questions. When does the optimization method work? How well does it work? We present results which are relatively easy to apply, and we show how to verify the technical details in practice with an example.

The remainder of the paper is organized as follows. In Section 2, we introduce a motivating example that will serve as the underlying thread of this paper. This example is a problem that is easy to describe, but difficult to solve when its underlying structure is high dimensional. We show how simulation optimization is relevant to the example. Then, in the next three sections of the paper, we discuss gradient-based techniques for solving the optimization problem (1). Section 3 gives an overview of gradient estimation techniques. Section 4 reviews stochastic approximation and Section 5 reviews sample average approximation. In Section 6 we describe the results of some limited experiments with the example of Section 2. Finally, in Section 7, we compare these two classes of methods and point out some potential difficulties in their use.

2 MOTIVATING EXAMPLE

Henceforth, all vectors are assumed to be column vectors, and x^T denotes the transpose of x.

Multi-dimensional Newsvendor Problem: Consider a firm that manufactures q products using p different resources. Suppose that a manager must decide on a non-negative resource vector $K \in \mathbb{R}^p_+$, before the product demand vector $D \in \mathbb{R}^q_+$ is observed. After the demand becomes known,

the manager chooses a production vector $x \in \mathbb{R}_+^q$ so as to maximize the operating profit in the following linear program:

$$\begin{aligned} \mathcal{P}_{(K,D)}: \quad \max_{x\in\mathbb{R}_+^q} \quad & v^T x \\ \text{s.t.} \quad & Ax \le K \text{ (capacity constraints)} \\ & x \le D \text{ (demand constraints).} \end{aligned}$$

Here, $v \in \mathbb{R}_+^q$ is a q-vector whose jth component represents the unit margin for product j (that is, revenue minus processing cost), and A is a $p \times q$ matrix whose component (i, j) represents the amount of resource i required to produce one unit of product j. Let $\pi(K,D)$ denote the maximal operating profit function for a given resource level K and a given demand D. This is precisely the optimal objective value of the problem $\mathcal{P}_{(\mathcal{K},\mathcal{D})}$. Then $\pi(K,D) = v^T x^*(K,D)$, where $x^*(K,D)$ is an associated optimal production vector.

Suppose that the demand D can be viewed as a random variable and the probability distribution of D is known. Let $\Pi(K)$ denote the expected maximal operating profit, where

$$\Pi(K) = E\pi(K,D),$$

for all $K \in \mathbb{R}_+^p$. The manager's objective is now to choose the resource level K so as to maximize the expected maximal operating profit. This leads to the following stochastic optimization problem:

$$\min_{K\in\mathbb{R}_+^p} -\Pi(K). \tag{2}$$

This problem is known as the *multi-dimensional newsvendor problem*. We do not consider specifically shortage penalties and holding costs, but they can be easily incorporated into the operating profit function (Van Mieghem and Rudi 2002). For simplicity, we focus our attention on the single-period newsvendor model, but the structure of the optimal policy in the single-period model can easily be extended to a dynamic setting under reasonable conditions (Harrison and Van Mieghem 1999).

It is much more difficult to find the optimal resource level for the multi-dimensional problem than for the single-dimensional problem (which actually has a closed-form solution). In the following sections, we illustrate several approaches for dealing with simulation optimization problems. We discuss how the simulation optimization methods can be used in practice and present technical details in the context of this example.

3 GRADIENT ESTIMATION

Consider the expected maximal operating profit function $\Pi(K) = E[\pi(K,D)]$. Assume that D is a continuous random vector that is finite with probability 1. We would like to determine conditions under which $\Pi(K)$ is differentiable and the interchange of the expectation and the gradient is valid, that is,

$$\nabla\Pi(K) = \nabla E[\pi(K,D)] = E[\nabla_K \pi(K,D)]. \tag{3}$$

This is called infinitesimal perturbation analysis (IPA). If IPA is valid, then it gives an unbiased estimator of $\nabla\Pi(K)$ from a single simulation run.

We can verify the validity of the IPA gradient estimate of $\Pi(\cdot)$ by using duality theory of linear programming. We follow the proof of Proposition 2 in Harrison and Van Mieghem (1999). Consider the dual problem of the linear program $\mathcal{P}_{(\mathcal{K},\mathcal{D})}$:

$$\begin{aligned} \mathcal{D}_{(K,D)}: \quad \min_{(\lambda,\mu)\in\mathbb{R}_+^{p+q}} \quad & K^T\lambda + D^T\mu \\ \text{s.t.} \quad & A^T\lambda + \mu \ge v. \end{aligned}$$

Since D is finite, the primal problem has a finite optimal solution. Hence, the optimal value of the primal problem is equal to that of the dual problem. Let $\lambda(K,D)$ denote the optimal shadow value of the capacity constraint in the primal problem $\mathcal{P}_{(\mathcal{K},\mathcal{D})}$. It can be easily shown that for any $K, K_0 \in \mathbb{R}_+^p$ and given D,

$$\pi(K,D) \le \pi(K_0,D) + \lambda(K_0,D)^T(K - K_0), \tag{4}$$

and hence $\lambda(\cdot,D)$ is a subgradient of $\pi(\cdot,D)$. From linear programming theory, we know that $\pi(\cdot,D)$ is concave for any fixed D, and hence is differentiable except on a set L of Lebesgue measure zero. Thus $\lambda(\cdot,D)$ is unique and $\nabla_K\pi(\cdot,D) = \lambda(\cdot,D)$ except on L. Taking the expectation in Equation (4) yields that $E\lambda(\cdot,D)$ is a subgradient of $\Pi(\cdot) = E\pi(\cdot,D)$. Since D is a continuous random variable, L has probability measure zero. So $E\lambda(K,D)$ is unique for all $K \in \mathbb{R}_+^p$ so that $\Pi(\cdot)$ is differentiable and $\nabla\Pi(\cdot) = E\lambda(\cdot,D) = E\nabla_K\pi(\cdot,D)$.

The concavity of the function $\pi(\cdot,D)$ and Equation (4) play important roles in the previous proof. The function $\pi(\cdot,D)$ is differentiable at a fixed K w.p. 1 so that the right hand side of Equation (3) makes sense. Note that $\pi(\cdot,D)$ does not have to be differentiable everywhere. Expectation with respect to a continuous random variable D smooths the function $\Pi(\cdot)$. For further detailed properties of $\Pi(\cdot)$ and a more general discussion on stochastic linear programming, see Birge and Louveaux (1997).

Under a set of assumptions, IPA can be applied to a more general form of the expected performance, one that does not have to be obtained from a linear program. The following theorem provides sufficient conditions for the interchange of the expectation and gradient to be valid for the objective function $f(\theta)$ in (1). The key condition is that the gradient of $f(\cdot,\xi)$ must be uniformly dominated by

an integrable function of ξ. Since each component of the gradient $\nabla f(\theta)$ can be dealt with separately, we assume without loss of generality that $p = 1$ for this theorem.

Theorem 1 *Let $\theta_0 \in \Upsilon$, where Υ is an open interval, and let H be a measurable set such that $P(\xi \in H) = 1$. Suppose that for every $z \in H$, there is a set $D(z)$, where $D(z)$ is at most countable, such that*

(i) *$\forall z \in H$, $f(\cdot, z)$ is continuous everywhere in Υ,*

(ii) *$\forall z \in H$, $f(\cdot, z)$ is differentiable everywhere in $\Upsilon \backslash D(z)$,*

(iii) *there exists a function $\phi : H \to [0, \infty)$ such that*

$$\sup_{\theta \in \Upsilon \backslash D(z)} |f'(\theta, z)| \leq \phi(z),$$

$\forall z \in H$ with $E\phi(\xi) < \infty$, and

(iv) *$f(\theta, \xi)$ is almost surely differentiable at $\theta = \theta_0$, i.e.,*

$$P(\xi \in \{z : f'(\theta_0, z) \text{ exists.}\}) = 1.$$

Then $f(\cdot)$ is differentiable at $\theta = \theta_0$, and

$$f'(\theta_0) = Ef'(\theta_0, \xi).$$

Further, if $f'(\cdot, z)$ is continuous all over Γ for each $z \in H$, then $f(\cdot)$ is continuously differentiable on Γ.

For a proof, see Proposition 1 in L'Ecuyer (1995).

IPA is usually highly efficient when it is valid (i.e., yields an unbiased gradient estimator). If the gradient $\nabla_\theta f(\cdot, \xi)$ is known analytically for a fixed ξ, the computational effort to compute the gradient of $f(\theta)$ can be significantly reduced. IPA has been successfully applied to a number of real-world problems. Unfortunately, there are many other problems where IPA is not valid.

Example 1 *Suppose that ξ is a random variable with a density, and that we wish to compute $P(\xi > \theta)$ and its derivative, for a given θ . Let $f(\theta) = P(\xi > \theta)$ and $f(\theta, \xi) = I(\xi > \theta)$, where I is the indicator function. Then $f'(\theta, \xi) = 0$ except when $\xi = \theta$, where it is undefined. But since ξ has a density, $P(\xi = \theta) = 0$ and $f'(\theta) < 0$ for some θ, so that $E[f'(\theta, \xi)] = 0 \neq f'(\theta)$. Thus the naïve IPA derivative estimator is not valid.*

In many cases, $f(\theta, \xi)$ can be replaced by a smoother alternative, and then IPA can be used. For example, one can use conditional expectation to smooth $f(\theta, \xi)$. This is called smoothed perturbation analysis (Fu and Hu 1997). For a more detailed overview of IPA and other perturbation analysis (PA) techniques, see Glasserman (1991), L'Ecuyer (1991) and Fu and Hu (1996).

If none of these techniques is available, or if they are too complicated to implement, we may appeal to finite difference (FD) schemes. Let $e_i = (0, ..0, 1, 0, ..., 0)$ denote the ith coordinate vector for $i = 1, ..., p$ and let $c > 0$. Then the ith component of the gradient $\nabla_\theta f(\theta)$ can be estimated by

$$\frac{f(\theta + ce_i, \xi^+) - f(\theta, \xi^-)}{c},$$

where ξ^+ and ξ^- are i.i.d. replications of ξ. This is the FD estimator using forward differences. If central differences are used, then

$$\frac{f(\theta + ce_i, \xi^+) - f(\theta - ce_i, \xi^-)}{2c}.$$

It is clear that the forward FD estimate requires the simulation of $p+1$ parameter values, whereas central FD requires $2p$ simulation replications. Thus, when $p > 1$, central FD involves more computational effort than forward FD. However, central FD usually has smaller bias than forward FD. The difference parameter c must be chosen carefully to achieve a balance between bias and variance. The bias of the FD estimators increases with c. But when c decreases to zero, the variance of FD estimators goes to infinity.

The variance of the FD estimators can be reduced by using common random numbers. Taking $\xi^+ = \xi^-$ in the above FD form yields forward and central FDC estimators (FD estimators with common random numbers). For small c, $f(\theta + ce_i, \xi^+)$ and $f(\theta - ce_i, \xi^+)$ are highly correlated, so a significant variance reduction can be obtained.

In practice, to estimate a gradient, one would take a sample of size n and use the corresponding sample average of a given gradient estimator. One might then be interested in the convergence rate of the mean square error (MSE) of the sample average. When IPA is valid and the variance is finite, under some uniform integrability condition, the MSE for IPA converges to zero at the canonical rate $O(n^{-1})$, from the central-limit theorem. On the other hand, in contrast to IPA, FD does not reach the canonical rate. Glynn (1989) and Zazanis and Suri (1993) obtained a subcanonical rate for FDC under a set of assumptions. But these assumptions turn out to be too loose when IPA applies. L'Ecuyer and Perron (1994) have shown that if the IPA estimator is unbiased and has finite variance, then the MSE of FDC is $O(n^{-1})$, the same as that of IPA, provided that there is one parameter (i.e., $p = 1$) and the difference parameter $c = c_n$ is of order $O(n^{-1/2})$. Thus when IPA estimation is too complicated, FDC can be a good alternative.

4 STOCHASTIC APPROXIMATION

Stochastic approximation (SA) methods are used to solve differentiable simulation optimization problems. They are analogous to the steepest-descent gradient search method in deterministic optimization, except that here the gradient does not have an analytic expression and must be estimated. Since the basic stochastic algorithms were introduced by

Robbins and Monro (1951) and Kiefer and Wolfowitz (1952), a huge amount of work has been devoted to this area.

The general form of the SA algorithm is a recursion where an approximation θ_n for the optimal solution is updated to θ_{n+1} using an estimator $g_n(\theta_n)$ of the gradient $\nabla f(\theta_n)$ of the objective function evaluated at θ_n. For a minimization problem, the recursion is of the form

$$\theta_{n+1} = \Pi_\Theta(\theta_n - a_n g_n(\theta_n)), \tag{5}$$

where Π_Θ denotes a projection of points outside Θ back into Θ, and $\{a_n\}$ is a sequence of positive real numbers such that

$$\sum_{n=1}^{\infty} a_n = \infty \text{ and } \sum_{n=1}^{\infty} a_n^2 < \infty. \tag{6}$$

The sequence $\{a_n\}$ is usually chosen to be of the form $a_n = a/n$ for all n, where a is a positive scalar, although other forms have their merits. When FD is used to obtain the gradient $g_n(\theta_n)$, the resulting procedure is called the Kiefer-Wolfowitz algorithm (Kiefer and Wolfowitz 1952). Fu (1990) and L'Ecuyer and Glynn (1994) studied the SA method with IPA gradient estimation, as applied to the optimization of the steady-state mean of a single-server queue.

In the presence of non-convexity, the SA algorithm may only converge to a local minimum. Theorem 2 below is an immediate specialization of [Theorem 2.1, p. 127] of Kushner and Yin (2003), which gives conditions under which θ_n converges to a local minimizer θ^* a.s. as $n \to \infty$. We need the following definitions.

A *box* $B \subset \mathbb{R}^p$ is a set of the form

$$B = \{x \in \mathbb{R}^p : a(i) \le x(i) \le b(i), i = 1, \ldots, p\}.$$

For $x \in B$, define the set $\mathcal{C}(x)$ as follows. For x in the interior of B, $\mathcal{C}(x) = \{0\}$. For x on the boundary of B, $\mathcal{C}(x)$ is the convex cone generated by the outward normals of the faces on which x lies. A *first-order critical point* x of a $\mathcal{C}^1$ function $f : B \to \mathbb{R}$ satisfies

$$-\nabla f(x) = z \text{ for some } z \in \mathcal{C}(x).$$

A first-order critical point is either a point where the gradient $\nabla f(x)$ is zero, or a point on the boundary of B where the gradient "points towards the interior of B". Let $S(f,B)$ be the set of first-order critical points of f in B. We define the distance from a point x to a set S to be

$$d(x,S) = \inf_{y \in S} \|x - y\|.$$

The projection $y = \Pi_B x$ is a pointwise projection defined by

$$y_i = \begin{cases} a(i) & \text{if } x(i) \le a(i), \\ x(i) & \text{if } a(i) < x(i) < b(i), \text{ and} \\ b(i) & \text{if } b(i) \le x(i) \end{cases}$$

for each $i = 1, \ldots, p$.

Let $(\mathcal{G}_n : n \ge 0)$ be a filtration, where the initial guess θ_0 is measurable with respect to $\mathcal{G}_0$ and $g_n(\theta_n)$ is measurable with respect to $\mathcal{G}_{n+1}$ for all n.

Theorem 2 *Let B be a box in $\mathbb{R}^p$ and $f : \mathbb{R}^p \to \mathbb{R}$ be $\mathcal{C}^1$. Suppose that for $n \ge 0$, $\theta_{n+1} = \Pi_B(\theta_n - a_n g_n(\theta_n))$ with the following additional conditions.*

- *(i)* *The conditions (6) hold.*
- *(ii)* $\sup_n E\|g_n(\theta_n)\|^2 < \infty$.
- *(iii)* *There is a sequence of random variables $\{\beta_n\}$ such that*

$$E[g_n(\theta_n)|\mathcal{G}_n] = \nabla f(\theta_n) + \beta_n, \text{ for all } n \ge 0,$$

where

$$\sum_{i=1}^{\infty} |a_i \beta_i| < \infty \text{ w.p.1.}$$

Then,

$$d(\theta_n, S(f,B)) \to 0$$

as $n \to \infty$ a.s. Moreover, suppose that $S(f,B)$ is a discrete set. Then, on almost all sample paths, θ_n converges to a unique point in $S(f,B)$ as $n \to \infty$.

Note that the limiting points in $S(f,B)$ can be random. If there exists a unique optimal solution, then θ_n converges to that point. Assumption (*iii*) implies that the bias sequence $\{\beta_n\}$ is asymptotically negligible. In the Kiefer-Wolfowitz algorithm, β_n represents the FD bias. In the classical SA methods, the function $g_n(\cdot)$ does not depend on n. In this case, a sufficient condition for the condition (ii) is that $\sup_{\theta \in \Theta} E||g_0(\theta)||^2 < \infty$. When the experimental design varies with n or variance reduction methods are used, we have the n-dependent function $g_n(\cdot)$.

Theorem 2 can be extended by considering a parameter set Θ, which is defined by a number of smooth constraints. Often Θ can be only estimated by noisy observations on ξ. For a discussion of how the SA algorithm can be used to solve constrained optimization problems with noisy constraints, see Kushner and Clark (1978) and Kushner and Yin (2003). It is known that choosing the gain sequence $\{a_n\}$ according to an analog of the deterministic Newton-Raphson algorithm provides an optimal form of the algorithm, that is, the estimator θ_n of θ^* converges at the highest possible rate $n^{-1/2}$. However, this procedure requires a priori knowledge of the Hessian matrix of $f(\theta)$, which is often difficult or

impossible to obtain in practice. Polyak and Juditsky (1992) demonstrate the asymptotic normality of the SA algorithm under milder conditions by averaging the estimates θ_n. See Kushner and Yin (2003) for more details about the asymptotic properties of the various SA algorithms.

Let us consider our motivating example, the multidimensional newsvendor problem. Suppose that each resource level has lower and upper bounds, i.e., $K \in B = \{K = (K(1),...,K(p)) \in \mathbb{R}^p : a(i) \le K(i) \le b(i), i = 1,..,p\}$, for some $0 < a(i) \le b(i) < \infty, i = 1,...,p$. Note that $\pi(\cdot, D)$ is concave on $\mathbb{R}^p_+$ for any fixed D, and thus $\Pi(\cdot)$ is concave. Since B is convex and compact, the problem

$$\min_{K \in B} -\Pi(K) \tag{7}$$

has a unique optimal solution K^*. The SA algorithm to search for the optimal K^* using an IPA estimator is as follows.

Stochastic Approximation

Initialization: Choose K_0.
For $n = 1$ to N_1
  Generate the i.i.d. sample $D_{n,i} \sim D$, $i = 1,...,N_2$, independent of all else.
  Compute

$$g_{n-1}(K_{n-1}) = \frac{1}{N_2} \sum_{i=1}^{N_2} \lambda(K_{n-1}, D_{n,i}), \text{ and}$$

$$K_n = \Pi_B(K_{n-1} + a_{n-1} g_{n-1}(K_{n-1})).$$

  If $\|a_{n-1} g_{n-1}(K_{n-1})\| < \varepsilon$ then exit loop.
Next n
Set $K^* = K_n$.

Once the estimate K_n of K^* is obtained, the expected profit $\Pi(K^*)$ can be estimated using $\hat{\Pi}_m = \frac{1}{m} \sum_{j=1}^{m} \pi(K_n, D_j)$, where $D_1,...,D_m$ is an i.i.d. sample independent of all the samples used for estimating K^*.

To verify the convergence of K_n, we will show that the IPA estimator in the above algorithm satisfies the conditions in Theorem 2. Since the IPA estimator is unbiased and $g_n(\cdot)$ does not depend on n in this case, it suffices to verify that $\sup_{K \in B} E||g_0(K)||^2 < \infty$. In fact, $\sup_{(K,D) \in B \times \mathbb{R}^q_+} ||\lambda(K,D)|| < \infty$. To see why, note that $\{x \in \mathbb{R}^q_+ : Ax \le K, K \in B\}$ is bounded. Then there exists some $U \in \mathbb{R}^q_+$ such that $\{x \in \mathbb{R}^q_+ : Ax \le K, K \in B\} \subset \Delta = \{x \in \mathbb{R}^q_+ : x \le U\}$. Then for any $D \in \mathbb{R}^q_+$, there exists some $\tilde{D} \in \Delta$ such that for any $K \in B$, $\pi(K,D) = \pi(K,\tilde{D})$, and hence $\lambda(K,D) = \lambda(K,\tilde{D})$. Therefore,

$$\{\lambda(K,D) : (K,D) \in B \times \mathbb{R}^q_+\}$$
$$= \{\lambda(K,D) : (K,D) \in B \times \Delta\}.$$

Since $\pi(\cdot,\cdot)$ is concave function on $\mathbb{R}^p_+ \times \mathbb{R}^q_+$ and for any D such that $||D||$ is small enough, $\pi(\cdot, D)$ is constant on B, $\sup_{(K,D) \in B \times \Delta} ||\lambda(K,D)||$ is bounded. Then clearly,

$$\sup_{K \in B} E||g_0(K)||^2 \le \sup_{(K,D) \in B \times \Delta} ||\lambda(K,D)||^2 < \infty.$$

The stochastic approximation procedure is easy to implement and exhibits good performance with appropriately chosen step sizes. But the algorithm is extremely sensitive to the choice of step size. Choosing good values is a non-trivial problem. Various procedures have been developed in which the step sizes are adaptively updated as the number of iterations grows. For more details, see Ruppert (1985).

5 SAMPLE AVERAGE APPROXIMATION

Another standard method to solve Problem (1) is that of sample average approximation (SAA). This method approximates the original simulation optimization problem (1) with a deterministic optimization problem. The framework is as follows: Let N be a positive integer and suppose that we generate the independent random sample $\xi_1, \ldots, \xi_N$. For a fixed θ, define the sample mean over $(f(\theta, \xi_i) : 1 \le i \le N)$ as

$$\bar{f}_N(\theta) = \frac{1}{N} \sum_{i=1}^{N} f(\theta, \xi_i).$$

The SAA problem corresponding to (1) is

$$\min_{\theta \in \Theta} \bar{f}_N(\theta), \tag{8}$$

i.e., we minimize the *sample* average. Once the sample is fixed, $\bar{f}_N(\theta)$ becomes deterministic. Its values and gradient can be computed for a given value of the parameter θ. Consequently, the SAA problem (8) becomes a deterministic optimization problem and one can solve it using any convenient optimization algorithm. The algorithm can exploit the IPA gradients, which are exact gradients of $\bar{f}_N(\theta)$.

Generally, the SAA problem (8) is close to the original problem (1) when N is large. By the strong law of large numbers(SLLN), we have that $\bar{f}_N(\theta)$ converges to $f(\theta)$ w.p. 1 as $N \to \infty$. The above construction can be extended to optimization problems with noisy constraints by approximating the constraint functions with sample averages (Bastin, Cirillo, and Toint 2007).

Various forms of this method have been used by different researchers. The *stochastic counterpart method* by Rubinstein and Shapiro (1993) uses likelihood ratios to obtain the approximate optimization problem. In Robinson (1996) this approach is called the *sample path method*. Plambeck et al. (1996) use a SAA method with IPA gradient

estimates to solve convex performance functions in stochastic systems, with extensive computational results. Healy and Schruben (1991) have also studied this method, which is called *retrospective optimization*. Chen and Schmeiser (2001) developed retrospective approximation algorithms for stochastic root finding problems. For an introduction to the SAA approach, the reader is referred to Shapiro (2003).

We next look at conditions under which we can prove that the SAA method will converge; that is, under which the optimal solutions of the SAA problem (8) approach the set of optimal solutions of the original problem as N grows. We give convergence results based on uniform convergence. Often, this is sufficient for practical applications. Let $\hat{v}_N$ and v^* denote the optimal objective values of the SAA problem (8) and the true problem (1) respectively. Throughout this section, we assume that Θ is a non-empty compact set.

Theorem 3 *Suppose that*

(i) *the set $S(f,\Theta)$ of optimal solutions of the true problem (1) is non-empty and contained in Θ,*
(ii) *the function $f(\cdot)$ is finite valued and continuous on Θ, and*
(iii) *$\bar{f}_N$ converges to f uniformly on Θ, that is*

$$\sup_{\theta\in\Theta}\left|\frac{1}{N}\sum_{i=1}^{N} f(\theta,\xi_i) - Ef(\theta,\xi)\right| \to 0$$

as $N\to\infty$ a.s.

Let $\hat{\theta}_N$ be an optimal solution of the SAA problem (8). Then $\hat{v}_N \to v^*$ and $d(\hat{\theta}_N, S(f,\Theta)) \to 0$ a.s. as $N\to\infty$.
For a proof, see Proposition 6 in Shapiro (2003).

The assumption (ii) in the above proposition is a uniform version of the strong law of large numbers (ULLN). The following result shows that, in the convex case, the pointwise LLN ensures that the ULLN holds on a compact set.

Proposition 4 *Suppose that for every $z\in H$,*

(i) *the function $f(\cdot,z)$ is convex on Θ,*
(ii) *the law of large numbers (LLN) holds pointwise, that is $\bar{f}_N(\theta)$ converges to $f(\theta)$ as $N\to\infty$ a.s., for any fixed $\theta\in\Theta$, and*
(iii) *$f(\cdot)$ is finite valued on a neighborhood of Θ.*

Then $\bar{f}_N$ converges to f uniformly on Θ, as $N\to\infty$ a.s.
For a proof, see Proposition 2 and Corollary 3 in Shapiro (2003).

In practice, we often have non-convex stochastic problems. The following Proposition provides relatively simple conditions for the ULLN to hold without the convexity assumption. We say that $f(\theta,\xi)$ is *dominated* by an integrable function $h(\cdot)$ if $Eh(\xi)<\infty$ and for every $\theta\in\Theta$, $|f(\theta,\xi)|\le f(\xi)$ a.s.

Proposition 5 *Suppose that for every $z\in H$,*

(i) *the function $f(\cdot,z)$ is continuous on Θ, and*
(ii) *$f(\theta,z)$ is dominated by an integrable function.*

Then $f(\theta)$ is finite valued and continuous on Θ and $\bar{f}_N$ converges to f uniformly on Θ, as $N\to\infty$ a.s.
For a proof, see Proposition 7 in Shapiro (2003).

Now let us return to the multidimensional newsvendor problem with bounded resource levels as explained in Section 4. To estimate the optimal expected profit $\Pi(K^*)$, we can consider the following sample average approximation algorithm.

Sample Average Approximation

Choose a positive integer $n\ge 2$.
Generate the i.i.d. sample $\tilde{D}_i\sim D$, $i=1,\ldots,n$, independent of all else.
For a fixed K, define

$$\bar{\Pi}_n(K) = \frac{1}{n}\sum_{i=1}^{n}\pi(K,\tilde{D}_i).$$

Find $\hat{K}_n$, an optimal solution for the problem

$$\min_{K\in B} -\bar{\Pi}_n(K).$$

The sample average approximation function $-\bar{\Pi}_n(K)$ is piece-wise linear and convex, but not smooth. However, the sample average approximation function becomes smooth as the sample size n grows, so in practice, one can choose sufficiently large n, and then apply an algorithm for optimization of smooth function to solve the sample average approximation problem using the IPA gradient estimator $-n^{-1}\sum_{i=1}^{n}\lambda(K,\tilde{D}_i)$. In our implementation we use a quasi-Newton method with a line-search.

Once the estimate $\hat{K}_n$ of K^* is obtained from the SAA algorithm, we then estimate the expected profit $\Pi(K^*)$ via the sample average $\hat{\Pi}_m(\hat{K}_n) = m^{-1}\sum_{j=1}^{m}\pi(\hat{K}_n,D_j)$, where the sample $D_1,\ldots,D_m$ is independent of the sample $\tilde{D}_1,\ldots,\tilde{D}_n$.

It is easy to verify that $\hat{K}_n$ converges to K^*. We only need to show the ULLN. Since $-\pi(\cdot,D)$ is convex on B and the convex function $-\Pi(\cdot)$ is bounded on a neighborhood of B, the ULLN is satisfied by Proposition 4.

Theorem 3 ensures that, if $\hat{\theta}_N$ solves the SAA problem (8) exactly, then $\hat{\theta}_N$ converges to the set of optimizers of the limit function f. Moreover, if the true problem (1) has a unique optimal solution θ^*, then $\hat{\theta}_N\to\theta^*$. However, in the non-convex case, the best that we can hope for from a computational point of view is that $\hat{\theta}_N$ is a first-order critical point for the SAA problem. We next prove the convergence of the first-order critical points to those of the true problem.

Theorem 6 *Suppose that Θ is convex, and*

(i) *for every $z\in H$, the function $f(\cdot,z)$ is continuously differentiable on a neighborhood of Θ, and*

(ii) *the gradient components* $\frac{\partial}{\partial \theta_i} f(\theta, z)(i = 1, ..., p)$ *are dominated by an integrable function.*

Let $\hat{\theta}_N \in S(\bar{f}_N, \Theta)$ *be the set of first-order critical points of* $\bar{f}_N$ *on* Θ. *Then* f *is continuously differentiable over* Θ *and* $d(\hat{\theta}_N, S(f, \Theta)) \to 0$ *as* $N \to \infty$ *a.s.*

For a proof, see Theorem 3.1 in Bastin, Cirillo, and Toint (2007).

Note that the above assumptions (i) and (ii) are sufficient conditions for a valid IPA gradient estimator (i.e., $\nabla_\theta \bar{f}_N(\theta)$ is an unbiased estimator of $\nabla_\theta f(\theta)$). These two assumptions imply that $\nabla_\theta \bar{f}_N(\theta)$ converges to $\nabla_\theta f(\theta)$ uniformly on Θ.

Theorem 6 shows that $\hat{\theta}_N$ converges to the set of first-order critical points of f as $N \to \infty$. This does not guarantee that the sequence $\{\hat{\theta}_m\}$ converges almost surely, as was the case for stochastic approximation. In general we cannot guarantee this because when there are multiple critical points, the particular critical point chosen depends, among other things, on the optimization algorithm that is used. Of course, a simple sufficient condition that ensures convergence is the existence of a unique first-order critical point. This condition is clearly difficult to verify in practice.

There exists a well-developed statistical inference of estimators derived by the SAA method. That inference is incorporated into validation analysis and error bounds for obtained solutions. If the true problem has a unique optimal solution $\hat{\theta}^*$, under a set of conditions an optimal solution θ_N for the SAA problem converges to $\hat{\theta}^*$ at a stochastic rate of $O_p(N^{-1/2})$ and the bias $E[\hat{v}_N] - v^* = E[f(\hat{\theta}_N)] - f(\theta^*)$ is of order $o(N^{-1/2})$ (Shapiro 1993, Shapiro 2003).

6 NUMERICAL RESULTS

In this section we examine the performance of the SA and SAA methods discussed in Section 4 and Section 5 on a multi-dimensional newsvendor problem with two products and two resources. Let $c \in \mathbb{R}^2_+$ be a vector whose ith component represents the unit investment cost for resource i. Incorporating the investment cost into the operating profit function $\Pi(\cdot)$ leads to the following optimization problem:

$$\max_{K \in B} V(K) = \Pi(K) - c'K. \tag{9}$$

We set $v = (3, 2)$ and $c = (1, 1)$.

We study the four cases given in Table 1. In the first two cases, each resource is dedicated to one product so that the optimization problem (9) is separable and its solution can be easily obtained using the inverse cumulative distribution functions of D_1 and D_2. The optimal resource level K^* and the maximal expected value V^* for Case 1 and 2 are given in Table 2. In Case 3 and 4, the second resource is dedicated to both product 1 and 2, and the the demands for product 1 and product 2 are correlated. In these cases, the optimal solution can only be found by solving the simultaneous equations using the multivariate demand distribution. In general, it is not easy to solve this problem analytically.

The values $\hat{K}_{SA}$ and $\hat{V}_{SA}$ are, respectively, the estimated optimal resource level and the expected maximal value obtained from the SA method. Similarly, $\hat{K}_{SAA}$ and $\hat{V}_{SAA}$ are the values obtained from the SAA method. In the SA algorithm, we took $N_2 = 100$ to compute the IPA gradient estimate in each step of the algorithm. For the SAA method, we took $n = 100$ samples and obtained $\hat{K}_{SAA}$ by applying a quasi-Newton method with a linesearch using IPA gradients to the sample average approximation problem $\min_{K \in B} -(\bar{\Pi}_n(K) - c'K)$. To ensure a fair comparison, we allocated equal amounts of CPU time to both algorithms. As an estimator of V^*, we used the sample average of $\pi(K) - c'K$, evaluated at $K = \hat{K}_{SA}$ and $\hat{K}_{SAA}$, over $m = 10,000$ replicates.

Table 1: The Four Cases for the Multi-dimensional Newsvendor Example

Case	A	D_1	D_2	ρ
1	$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$	gamma(2,5)	gamma(2,10)	0
2		gamma(2,5)	unif(0,20)	0
3	$\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$	gamma(2,5)	gamma(2,10)	.5
4		gamma(2,5)	gamma(2,10)	.7

Table 2: Optimal Solutions and Values for the Multi-dimensional Newsvendor Example

Case	K^*	V^*
1	(11.45,16.78)	21.52
2	(11.45,10)	17.03

Table 3 shows that in all four cases, the estimated maximal value obtained from the SAA method is bigger than the one given by the SA method. This implies that the solutions obtained from the SAA procedure are closer to the true solutions than those from the SA method. In fact, Table 2 shows that in Case 1 and 2, $\hat{K}_{SAA}$ is fairly close to K^*. As we said earlier, the SA estimator is very sensitive to the step size parameters an. When we used this method, we tuned the parameters heuristically until reasonable performance was observed. A contour plot of the expected value surface for Case 4 appears in Figure 1. We see that the function is concave and close to a smooth function, and it has a maximum value near the point $(7, 22)$, which is close to the solution obtained from the SAA method.

7 CONCLUSIONS

We have provided a review of two gradient-based techniques for simulation optimization. It remains to compare the

Table 3: Simulation Results for the Multi-dimensional Newsvendor Example

Case	$\hat{K}_{SA}$	$\hat{V}_{SA}$	$\hat{K}_{SAA}$	$\hat{V}_{SAA}$
1	(9.12, 7.86)	18.43	(11.20,16.23)	21.57
2	(9.63,6.70)	16.35	(11.20,9.64)	17.08
3	(2.35, 10.53)	9.16	(7.15,21.54)	13.40
4	(1.64, 8.64)	7.78	(6.76,21.87)	13.22

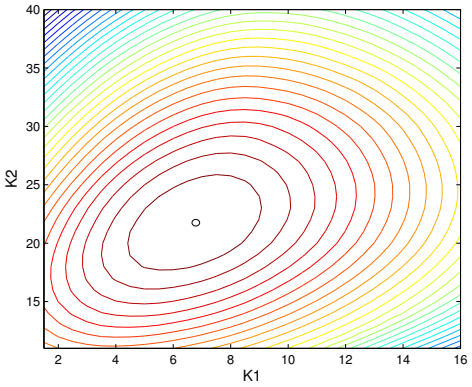

Figure 1: Contour Plot of $V(\cdot)$ for Case 4 with Runlength 500

stochastic approximation and sample average approximation methods, and to discuss certain performance issues that come up when these methods are used in practice.

The stochastic approximation method has nice asymptotic properties. We have seen that the stochastic approximation estimator approaches a first-order critical point for the original problem in the long term. However, in practice it requires many iterations to achieve the convergence. The stochastic approximation scheme is easy to implement and requires low computational effort per iteration. However, it is highly sensitive to the choice of certain tuning parameters (i.e., the step sizes). In order to achieve satisfactory performance, the method typically requires some adaptive modification of the tuning parameters.

The sample average approximation method substitutes a deterministic optimization problem instead of the original problem. It does not require any tuning parameters, so it is more robust than the stochastic approximation method. If a fast deterministic optimization algorithm can be used, then this method can be very effective on problems with many variables and/or complicated constraints. However, it can involve a large amount of computational effort.

Both methods work well when the underlying problem is well-structured: for example, when we are dealing with a smooth objective function, or when some relevant random functions satisfy certain moment conditions. But how can we identify the structure of the problem? As we have seen in our example, sample path analysis can be a clean way to analyze the problem structure. Unfortunately, this approach does not always work. It would be of great interest to obtain general results that could guarantee a "good" structure for whole classes of simulation optimization problems that arise in practice.

ACKNOWLEDGMENTS

The author thanks Shane Henderson for helpful comments and suggestions that substantially improved this paper.

REFERENCES

Andradóttir, S. 1998. A review of simulation optimization techniques. In *Proceedings of the 1998 Winter Simulation Conference*, ed. D. Medeiros, E. Watson, J. S. Carson, and M. S. Manivannan, 151–158.

Azadivar, F. 1999. Simulation optimization methodologies. In *Proceedings of the 1999 Winter Simulation Conference*, ed. P. A. Farrington, H. Black Nembhard, D. T. Sturrock, and G. W. Evans, 93–100.

Bastin, F., C. Cirillo, and P. L. Toint. 2007. Convergence theory for nonconvex stochastic programming with an application to mixed logit. *Mathematical Programming* 108:207–234.

Birge, J. R., and F. Louveaux. 1997. *Introduction to stochastic programming*. New York: Springer.

Chen, H., and B. W. Schmeiser. 2001. Stochastic root finding via retrospective approximation. *IIE Transactions* 33:259–275.

Fu, M. C. 1990. Convergence of the GI/G/1 queue using infinitesimal perturbation analysis. *Journal of Optimization Theory and Applications* 65:149–160.

Fu, M. C. 1994. Optimization via simulation: a review. *Annals of Operations Research* 53:199–247.

Fu, M. C. 2002. Optimization for simulation: theory vs. practice. *INFORMS Journal on Computing* 14:192–215.

Fu, M. C., F. Glover, and J. April. 2005. Simulation optimization: A review, new developments and applications. In *Proceedings of the 2005 Winter Simulation Conference*, ed. M. E. Kuhl, N. M. Steiger, F. B. Armstrong, and J. A. Joines, 83–95.

Fu, M. C., and J.-Q. Hu. 1996. A comparison of perturbation analysis techniques. In *Proceedings of the 1996 Winter Simulation Conference*, ed. J. M. Charnes, D. J. Morrice, D. T. Brunner, and J. J. Swain, 295–301.

Fu, M. C., and J.-Q. Hu. 1997. *Conditional Monte Carlo: Gradient estimation and optimization applications*. Boston: Kluwer.

Glasserman, P. 1991. *Gradient estimation via perturbation analysis*. The Netherlands: Kluwer.

Glynn, P. W. 1989. Optimization of stochastic systems via simulation. In *Proceedings of the 1989 Winter Simulation Conference*, 90–105.

Harrison, J. M., and J. A. Van Mieghem. 1999. Multi-resource investment stragies: Operational hedging under demand uncertainty. *European Journal of Operational Research* 113:17–29.

Healy, K., and L. W. Schruben. 1991. Retrospective simulation response optimization. In *Proceedings of the 1991 Winter Simulation Conference*, ed. B. L. Nelson, W. D. Kelton, and G. Clark, 901–907.

Henderson, S. G., and B. L. Nelson. (Eds.) 2006. *Handbook of simulation*. Elsevier. Forthcoming.

Kiefer, J., and J. Wolfowitz. 1952. Stochastic estimation of the maximum of a regression function. *Annals of Mathematical Statistics* 23:462–466.

Kushner, H. J., and D. C. Clark. 1978. *Stochastic approximation methods for constrained and unconstrained systems*. New York: Springer-Verlag.

Kushner, H. J., and G. G. Yin. 2003. *Stochastic approximation and recursive algorithms and applications*. 2nd ed. New York: Springer-Verlag.

L'Ecuyer, P. 1991. A unified view of the IPA, SF and LR gradient estimation techniques. *Management Science* 36:1364–1383.

L'Ecuyer, P. 1995. On the interchange of derivative and expectation for likelihood ratio derivative estimators. *Management Science* 41:738–748.

L'Ecuyer, P., and P. W. Glynn. 1994. Stochastic optimization by simulation: Convergence proofs for the GI/G/1 queue in steady-state. *Management Science* 40:1562–1578.

L'Ecuyer, P., and G. Perron. 1994. On the convergence rates of IPA and FDC derivative estimators. *Operations Research* 42(4):643–656.

Plambeck, E. L., B.-R. Fu, S. M. Robinson, and R. Suri. 1996. Sample-path optimization of convex stochastic performance functions. *Mathematical Programming* 75:137–176.

Polyak, B., and A. Juditsky. 1992. Acceleration of stochastic approximation by averaging. *SIAM Journals on Control and Optimization* 30(4):838–855.

Robbins, H., and S. Monro. 1951. A stochastic approximation method. *Annals of Mathematical Statistics* 22:400–407.

Robinson, S. M. 1996. Analysis of sample-path optimization. *Mathematics of Operations Research* 21:513–528.

Rubinstein, R. Y., and A. Shapiro. 1993. *Discrete event systems: Sensitivity analysis and stochastic optimization by the score function method*. Chichester: Wiley.

Ruppert, D. 1985. A Newton-Raphson version of the multivariate Robbins-Monro procedure. *Annals of Statistics* 13:236–245.

Shapiro, A. 1993. Asymptotic behavior of optimal solutions in stochastic programming. *Mathematics of Operations Research* 18:829–845.

Shapiro, A. 2003. Monte Carlo sampling methods. In *Stochastic Programming*, ed. A. Ruszczynski and A. Shapiro, Handbooks in Operations Research and Management Science. Elsevier.

Van Mieghem, J. A., and N. Rudi. 2002. Newsvendor networks: Inventory management and capacity investment with discretionarty activities. *Manufacturing and Service Operations Management* 4(4):313–335.

Zazanis, M. A., and R. Suri. 1993. Convergence rates of finite-difference sensitivity estimates for stochastic systems. *Operations Research* 41:694–703.

AUTHOR BIOGRAPHY

SUJIN KIM is a Visiting Assistant Professor in the School of Industrial Engineering at Purdue University. She received a Ph.D. in Operations Research from Cornell University in 2006. Her research interests are in simulation optimization and efficiency improvement techniques in Monte Carlo simulation. Her e-mail address is <sujin@purdue.edu>.