Proceedings of the 1986 Winter Simulation Conference
J. Wilson, J. Henriksen, S. Roberts (eds.)

OPTIMIZATION OF STOCHASTIC SYSTEMS

Peter W. Glynn
Department of Industrial Engineering
and Mathematics Research Center
University of Wisconsin-Madison
Madison, WI 53705, U.S.A.

ABSTRACT

This paper gives a short survey of Monte Carlo algorithms for stochastic optimization. Both discrete and continuous parameter stochastic optimization are discussed, with emphasis on the analysis of convergence rate. Some future research directions for the area are also indicated.

1. INTRODUCTION

Stochastic optimization is concerned with the general problem of optimization under uncertainty. This problem arises in contexts as diverse as long-range economic planning, warehouse inventory operations, automated manufacturing systems, and computer communications networks. In this tutorial, we hope to provide a hierarchal framework for stochastic optimization. Within this framwork, we will discuss the major research problems and indicate the current status of solution methodologies to these problems.

The scope of the survey given here is not intended to be exhaustive, nor do we intend to describe solution algorithms in great detail. Rather, we wish to give the reader of this brief paper a flavor for the current state of the art and future directions for research in the area. The opinions expressed here are subjective, and should be viewed as such. Having stated this, let us proceed.

2. A HIERARCHAL FRAMEWORK FOR STOCHASTIC OPTIMIZATION

Roughly speaking, stochastic optimization can be viewed in terms of the tree structure expressed by Figure 1. As a first cut, one can differentiate stochastic optimization according to whether the dimensionality of the space of decision variables is finite or infinite.

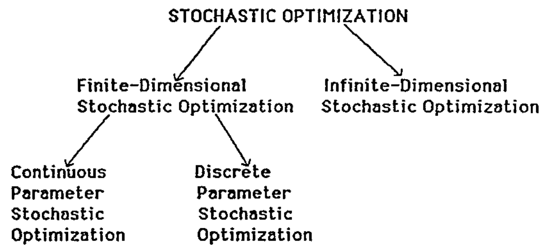
STOCHASTIC OPTIMIZATION
Finite-Dimensional Stochastic Optimization
Infinite-Dimensional Stochastic Optimization
Continuous Parameter Stochastic Optimization
Discrete Parameter Stochastic Optimization

FIGURE 1

i.) <u>infinite-dimensional stochastic optimization</u>: Here, one is usually concerned with optimization over a class of permissible control policies. This is typical of applications in which the objective is to determine a time-varying control policy $U = \{U(t,\omega) : t > 0\}$ which minimizes the expected cost of running a stochastic system: i.e. Find U^* to solve

$$\min_U EC(U(\cdot,\omega),\omega) \quad .$$

This problem of <u>stochastic control</u> can frequently be treated by dynamic programming arguments. In particular, it is often possible to obtain U^* by solving <u>Bellman's optimality equation</u>; see Bellman and Dreyfus (1962).

Of course, for complex systems, it is usually not possible to analytically solve for U^*. In such cases, it may be possible to find U^* by using numerical algorithms (i.e. value iteration or policy iteration) based on the theory of Markov decision processes; see, for example, Denardo (1982). Even if U^* can not be determined numerically, the theory of dynamic programming may be useful in obtaining a qualitative characterization of the optimal control. For example, it is known, under quite general conditions, that the optimal control for operation of certain inventory systems is characterized by two critical numbers s and S (a re-order point and an ordering amount); these are the so-called (s,S) inventory systems. Thus, dynamic programming has reduced the optimization problem from that of determining the infinite-dimensional control U^* to solving for the two parameters s and S (i.e. a two-dimensional optimization problem).

ii.) finite-dimensional stochastic optimization: Whereas the infinite-dimensional problem generally involves optimization over some infinite-dimensional function class, finite-dimensional optimization concerns optimization (roughly speaking) over some subset of Euclidian space. To be precise, let (Ω, F, P_θ) be a probability space in which the parameter θ takes values in $\Lambda \subseteq R^d (d \geq 1)$. The probability measure P_θ describes how the random environment is affected by the choice of θ. For each $\theta \in \Lambda$, let $X(\theta)$ be a real-valued random variable corresponding to the cost of running the "system" under θ. Then,

$$\alpha(\theta) = \int_\Omega X(\theta,\omega)P_\theta(d\omega)$$

is the expected cost of running the system under θ. The general finite-dimensional stochastic optimization problem involves finding $\theta^* \in \Lambda$ to minimize $\alpha(\theta)$ subject (possibly) to constraints of the form

$$\int_\Omega Y_i(\theta,\omega)P_\theta(d\omega) \geq 0 \text{ ,}$$

$1 \leq i \leq m$, where $\{Y_i(\theta) : 1 \leq i \leq m\}$ is a collection of "random constraints".

If the above expectations can be evaluated analytically in closed form, then the problem is amenable to solution by standard non-linear programming algorithms. Otherwise, any numerical procedure for solving this problem must (at least implicitly) numerically integrate in order to calculate the expectations. If the structure of the numerical integration is complicated, then Monte Carlo simulation may be the only viable procedure for performing the required integrations.

The remainder of this paper is therefore devoted to Monte Carlo algorithms for solving complex finite-dimensional stochastic optimization problems. Before proceeding, it is worth noting that the best possible convergence rate possible with "pure" Monte Carlo algorithms is of order $t^{-1/2}$ in the computational effort t. This rate of convergence is essentially a consequence of the "central limit theorem" behavior that is typical of good Monte Carlo estimators $\alpha(t)$ for a parameter α:

$$t^{1/2}(\alpha(t) - \alpha) \Longrightarrow \sigma N(0,1)$$

as $t \to \infty$. Note that a convergence rate of order $t^{-1/2}$ implies that one must multiply the run length by a factor of 100 to obtain an additional significant figure of accuracy, a factor of 10000 to obtain two additional significant figures, etc. Thus, even the best possible Monte Carlo convergence rate is extremely slow. As a result, one should never use Monte Carlo simulation to solve a problem if conventional numerical approaches can be easily applied to the problem at hand.

Returning to the optimization context under study here, it turns out that some Monte Carlo algorithms converge even more slowly than at rate $t^{-1/2}$; these methods should clearly be avoided, if at all possible. We will come back to this point later in the paper.

As for the finite-dimensional stochastic optimization problem currently under discussion, there are two basic variants:

a.) Λ discrete

b.) Λ continuous.

As in deterministic mathematical programming, the techniques used to solve the discrete and continuous problems are very different. As in the deterministic case, the continuous optimization problem is, in some sense, easier than the discrete problem. It is to be expected that solution methodologies for the discrete problem will need, for the most part, to be tailor-made to the application. Continuous algorithms will be more robust, in the sense that they can be successfully applied to fairly general classes of problems.

3. CONTINUOUS PARAMETER STOCHASTIC OPTIMIZATION

Throughout this section, we shall assume that we are working with the unconstrained problem in which $\Lambda = R^d$ and $Y_i(\theta) = 0$ for all θ; for constrained algorithms, we suggest looking at Kushner and Clark (1978).

Roughly speaking, there are four major classes of Monte Carlo algorithms currently available for continuous parameter unconstrained optimization; see Figure 2.

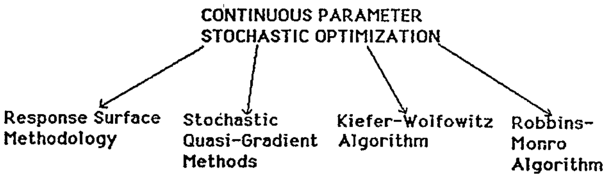
CONTINUOUS PARAMETER
STOCHASTIC OPTIMIZATION
Response Surface Methodology
Stochastic Quasi-Gradient Methods
Kiefer-Wolfowitz Algorithm
Robbins-Monro Algorithm

FIGURE 2

i.) <u>response surface methodology</u>: The basic idea here is to select n points $\theta_1, \theta_2, \ldots, \theta_n$ in the decision variable space. One then estimates $\alpha(\theta_i)$, $1 < i < n$ by $\hat{\alpha}_i(\theta_i)$ and fits a "smooth" surface to the n pairs $(\theta_i, \hat{\alpha}_i(\theta_i))$, $1 < i < n$. The "smooth" surface usually consists of some polynomial approximation. A solution for the original optimization problem is then obtained by optimizing the fitted surface (i.e. the optimizer of the fitted surface is taken to be the optimizer of the original problem). This method can give poor results if:

a.) n is not large enough

b.) the points θ_i are not selected appropriately

c.) the estimators $\hat{\alpha}_i(\theta_i)$ are not "close enough" to $\alpha_i(\theta_i)$

d.) the original surface $\alpha(\cdot)$ is not well described by a polynomial approximation.

ii.) <u>stochastic quasi-gradient methods</u>: The idea here is to adapt gradient-driven deterministic mathematical programming algorithms to the stochastic context by substituting Monte Carlo finite-difference estimates for the gradients; see Polyak (1976) and Ermoliev (1983). A main difficulty with this approach is that Monte Carlo finite-difference gradient estimation tends to have a <u>very</u> slow rate of convergence. To see this, let

$$\Delta\hat{\alpha}_n(\theta;h) = h^{-1}[\hat{\alpha}_n^1(\theta+h) - \hat{\alpha}_n^2(\theta)]$$

where $\hat{\alpha}_n^1$ and $\hat{\alpha}_n^2$ are independently generated sample means of $X(\theta+h)$ (under $P_{\theta+h}$) and $X(\theta)$ (under P_θ), respectively. Then,

$$\text{MSE}(\Delta\hat{\alpha}_n(\theta;h)) = E(\Delta\hat{\alpha}_n(\theta;h) - \alpha'(\theta))^2$$

$$= \text{var}(\Delta\hat{\alpha}_n(\theta;h)) + [E\Delta\hat{\alpha}_n(\theta;h) - \alpha'(\theta)]^2 .$$

Expanding $\alpha(\cdot)$ via a Taylor series, one finds that for h small,

$$E\Delta\hat{\alpha}_n(\theta;h) \approx \alpha'(\theta) + h\alpha''(\theta)/2$$

so that

$$\mathrm{MSE}(\Delta\hat{\alpha}_n(\theta;h)) \approx 2h^{-2}\mathrm{var}_\theta X(\theta)/n + h^2\alpha''(\theta)^2/4 .$$

If one chooses h to depend on n in the optimal way (in reality, one can't generally do this well), then

$$h = h_n \propto n^{-1/4}$$

and

$$\mathrm{MSE}(\Delta\hat{\alpha}_n(\theta;h_n)) \propto n^{-1/2} .$$

Consequently, the rate at which $\Delta\hat{\alpha}_n(\theta;h_n)$ converges to $\alpha'(\theta)$ is $n^{-1/4}$, in the number of observations n generated. (Observe that the root mean square error converges at rate $n^{-1/4}$; for more on this issue, see Fox and Glynn (1986) and Zazanis and Suri (1985). It turns out that if central differences, rather than finite differences, are used, then one can obtain faster rates of convergence, although the rate is still slower than $n^{-1/2}$ (the optimal Monte Carlo rate).

In any case, this analysis suggests that any algorithm which attempts to consistently estimate the gradient via Monte Carlo finite differences will converge at a rate slower than $t^{-1/2}$ in the computational effort t.

iii.) Kiefer-Wolfowitz algorithm: To simplify the discussion, assume $d = 1$. If $\alpha(\cdot)$ is differentiable, an optimizer θ^* must be a root of

$$\alpha'(\theta^*) = 0 ,$$

furthermore, if θ^* is to be a minimizer, $\alpha'(\theta^*+) > 0$ and $\alpha'(\theta^*-) < 0$. To find a root of α', consider Newton's method:

$$\theta_{n+1} = \theta_n - \alpha'(\theta_n)/\alpha''(\theta_n) .$$

Assume that $\alpha''(\theta^*) > 0$. For large n, $\alpha''(\theta_n) \approx \alpha''(\theta^*)$, so the algorithm looks like

$$\theta_{n+1} \approx \theta_n - \alpha'(\theta_n)/\beta$$

where $\beta = \alpha''(\theta^*)$. Since we cannot evaluate α or α' in closed form, we use a Monte Carlo estimate instead. Suppose that $Y_n(\theta_n)$ is generated so that

$$E\{Y_n(\theta_n) \mid \theta_1,\ldots,\theta_n\} \approx \alpha'(\theta_n) .$$

This suggests choosing θ_{n+1} so that

$$\theta_{n+1} \approx \theta_n - Y_n(\theta_n)/\beta .$$

However, $Y_n(\theta_n)$ may be a very "noisy" estimate of $\alpha'(\theta_n)$; consequently, to "damp out" the random effects, weight the "current" value θ_n by its corresponding sample size n, and the proposed "new" point $\theta_n - Y_n(\theta_n)/\beta$ by its corresponding sample size 1:

$$\theta_{n+1} \approx (\tfrac{n}{n+1})\theta_n + (\tfrac{1}{n+1})(\theta_n - Y_n(\theta_n)/\beta)$$
$$= \theta_n - (n+1)^{-1}Y_n(\theta_n)/\beta .$$

The Kiefer-Wolfowitz (KW) algorithm involves using

$$Y_n(\theta_n) = c_n^{-1}(X^1(\theta_n+c_n) - X^2(\theta_n))$$

when $X^1(\theta_n+c_n)$ and $X^2(\theta_n)$ are independently generated from $P_{\theta_n+c_n}$ and P_{θ_n}, respectively.

Note that the KW algorithm does not attempt to consistently estimate $\alpha'(\theta_n)$ before iterating to θ_{n+1}. However, because of the use of finite differences to evaluate $\alpha'(\theta_n)$, the method suffers from a slow convergence rate. When the KW algorithm converges to θ^*, the convergence rate, except in rare cases, is $t^{-1/3}$ in the computational effort t; see Sacks (1958).

iv.) Robbins-Monro algorithm: The Robbins-

Monro (RM) algorithm is obtained when one can find an unbiased estimator $Y_n(\theta_n)$ for $\alpha'(\theta_n)$, in which case

$$\theta_{n+1} = \theta_n - (n+1)^{-1}Y_n(\theta_n)/\beta \ ;$$

since the algorithm is driven by unbiased derivative estimates, the convergence rate (except in rare cases) is $t^{-1/2}$, in the computational effort t. In other words, RM algorithms attain the best possible Monte Carlo convergence rate.

More precisely, the convergence rate is of the form $\sigma t^{-1/2}$. To make σ as small as possible, choosing $\beta = \alpha''(\theta^*)$ is optimal. However, the algorithm converges with a $t^{-1/2}$ convergence rate regardless of the choice of β.

Given that RM algorithms achieve the best possible rate of convergence, it is incumbent upon the user to find an unbiased estimator for the derivative of α. Recall that

$$\alpha(\theta) = E_\theta X(\theta) \ .$$

If the expectation did not depend on θ (ie. if $\alpha(\theta) = EX(\theta)$), then (assuming the interchange of derivative and expectation was valid), it would follow that $\alpha'(\theta) = EX'(\theta)$. Thus, by driving the RM algorithm according to the iteration

$$\theta_{n+1} = \theta_n - (n+1)^{-1}X'(\theta_n)/\beta \ ,$$

we would obtain a convergence algorithm. However, P_θ does depend on θ.

To eliminate this dependence, two basic approaches are possible. Both are most easily illustrated when $\Omega = R^1$, in which case P_θ is determined by a distribution function $F(\theta,\cdot)$. The first idea involves observing that

$$\begin{aligned}\alpha(\theta) &= \int_{-\infty}^{\infty} X(\theta,\omega)F(\theta,d\omega)\\ &= \int_0^1 X(\theta,F^{-1}(\theta,\omega))d\omega\\ &= EX(\theta,F^{-1}(\theta,U))\end{aligned}$$

where $F^{-1}(\theta,\cdot)$ is the inverse distribution function defined by $F^{-1}(\theta,x) = \sup\{y : F(\theta,y) < x\}$, and U is a uniform $(0,1)$ r.v. Thus, if $X(\cdot,\cdot)$ and $F^{-1}(\cdot,x)$ are differentiable, it follows that

$$\alpha'(\theta) = EW(\theta)$$

where

$$W(\theta) = \frac{\partial X}{\partial x_1}(\theta,F^{-1}(\theta,U)) + \frac{\partial X}{\partial x_2}(\theta,F^{-1}(\theta,U)) \cdot \frac{d}{d\theta}F^{-1}(\theta,U) \ .$$

If we generate the sequence $\{\theta_n : n > 1\}$ via the iteration

$$\theta_{n+1} = \theta_n - (n+1)^{-1}W(\theta_n)/\beta \tag{3.1}$$

($W(\theta_n)$ generated independently of the past history), then the algorithm is in the setting of the classical RM algorithm, and we may expect $n^{-1/2}$ convergence.

A second way of eliminating dependence of P_θ upon θ goes as follows. Assume that $F(\theta,\cdot)$ has a density $f(\theta,\cdot)$. Let $g(\cdot)$ be the density of a r.v. Z for which the density is positive everywhere (eg. $Z \sim N(0,1)$). Then

$$\begin{aligned}\alpha(\theta) &= \int_{-\infty}^{\infty} X(\theta,x)f(\theta,x)dx\\ &= \int_{-\infty}^{\infty} X(\theta,x)\frac{f(\theta,x)}{g(x)} \cdot g(x)dx\\ &= EX(\theta,Z)f(\theta,Z)/g(Z)\end{aligned}$$

where Z has density g. If $X(\cdot,x)$ and $f(\cdot,x)$ are differentiable, then one expects that

$$\alpha'(\theta) = EW(\theta)$$

where

$$W(\theta) = \frac{d}{d\theta} X(\theta,Z) \cdot \frac{f(\theta,Z)}{g(Z)} + \frac{X(\theta,Z)}{g(Z)} \cdot \frac{d}{d\theta} f(\theta,Z) . \tag{3.2}$$

Again we may expect that (3.1) yields $n^{-1/2}$ convergence.

Note that the first method involves representing the response in terms of a single random variable; thus, the first method can be viewed as an application of common random numbers to derivative estimation. (See Rubinstein (1981).) It turns out that this approach applies also to discrete-event systems; the estimator $W(\theta)$ developed in the discrete-event context is the perturbation analysis estimator due to Ho and Suri (see Suri (1983)).

The second technique is an application of importance sampling to derivative estimation (see [11]). For a version of $W(\theta)$ applicable to Markov chains, see Glynn (1986). For those familiar with statistics, one may interpret the ratio $f(\theta,Z)/g(Z)$ as a likelihood ratio; as a consequence, we refer to Monte Carlo derivative estimators based on (3.2) as likelihood ratio derivative estimators.

It should be noted that the common random number approach requires that X be smooth in both arguments, whereas the likelihood ratio technique postulates smoothness only in the θ-component. As a result, the common random number approach is more susceptible to difficulties in interchanging the derivative and expectation; see Heidelberger (1986).

It should be evident that a major element of continuous optimization concerns evaluation of the gradient of α. The gradient can be viewed as an indication of the sensitivity of α to perturbations in the vector θ. Recently, Schruben (1986) has investigated a frequency domain approach to obtaining sensitivity estimates; it is to be anticipated that such an approach will prove useful in the optimization process.

To conclude this section, we note that the best that we can generally hope to achieve with an iterative algorithm of the type described above is moderately fast (i.e. order $t^{-1/2}$) convergence to a local optimizer. As in the deterministic mathematical programming context, one should never expect convergence to the global optimizer. Of course, if the global optimizer is the unique local optimizer, then one can expect convergence (as in the deterministic case).

4. DISCRETE PARAMETER STOCHASTIC OPTIMIZATION

As in Section 3, we assume that $Y_i(\theta) = 0$ for all θ. However, Λ is now hypothesized to be a discrete subset of $\mathbf{R}^d$. Note that the problems indexed by Λ need have no relation to each other. This is in contrast to the continuous setting where continuity and differentiability considerations force the problems to "mesh" together.

The unstructured discrete optimization problem takes the form:

$$\min_{\theta_i \in \Lambda} \alpha(\theta_i) . \tag{4.1}$$

Of course, (4.1) can be viewed as the "selection of best system" problem. The enormous statistical literature on ranking and selection applies to this setting. Note, however, that the formulation (4.1) and the associated selection algorithms make essentially no use of problem structure that may be present. To some extent, problem structure can be used by applying specially tailored variance reduction techniques to estimating $\alpha(\theta_i)$, $\theta_i \in \Lambda$.

A more non-trivial application of problem structure is presented in Glynn and Sanders (1986). The idea there is to deform the discrete optimization problem into one that has an easily determined solution. For example, the pallet loading problem for an automatic assembly system has a trivial solution if the machine jam rate λ is zero. One now increases λ to the "real world" jam rate λ^*, following the optimizer through λ space to the λ^*-optimizer.

This type of approach has met with considerable success in the deterministic mathematical programming context (see Allgower and Georg (1983)). Appropriate Monte Carlo algorithms for following the optimal path are currently under investigation.

To summarize, discrete parameter Monte Carlo optimization is in its infancy, and is not as well developed as in the continuous case. As in the deterministic setting, one expects efficient Monte Carlo algorithms to exploit problem structure as fully as possible. For example, in buffer sizing problems, it should be possible to use the nested nature of the k-buffer problem within the (k+1)-buffer problem.

ACKNOWLEDGEMENTS

The preparation of this paper was supported by the National Science Foundation Grant ECS-840-4809 and sponsored by the United States Army under Contract No. DAAG29-80-C-0041.

REFERENCES

[1] Allgower, E. L. and Georg, K. (1983). Predictor-corrector and simplicial methods for approximating fixed points and zero points of nonlinear mappings. Mathematical Programming - The State of the Art - Bonn 1982. Springer-Verlag, New York.

[2] Bellman, R. E. and Dreyfus, S. E. (1962). Applied Dynamic Programming. Princeton University Press, Princeton, NJ.

[3] Denardo, E. (1982). Dynamic Programming. Prentice-Hall, Englewood Cliffs, NJ.

[4] Ermoliev, Y. (1983). Stochastic quasi-gradient methods and their application to system optimization. Stochastics 9, 1-36.

[5] Fox, B. L. and Glynn, P. W. (1986). Replication schemes for limiting expectations. Technical Report, University of Montreal.

[6] Glynn, P. W. (1986). Stochastic approximation for Monte Carlo simulation. To appear elsewhere in this volume.

[7] Glynn, P. W. and Sanders, J. L. (1986). Monte Carlo optimization of stochastic systems: Two new approaches. Proceedings of the 1986 ASME Computers in Engineering Conference.

[8] Heidelberger, P. (1986). Limitations of infinitesimal perturbation analysis. Research Report, IBM, Yorktown Heights, NY.

[9] Kushner, H. J. and Clark, D. S. (1978). Stochastic Approximation Methods for Constrained and Unconstrained Systems. Springer-Verlag, New York.

[10] Polyak, B. T. (1976). Convergence and convergence rate of iterative stochastic algorithms I. General case. Automatika i Telemekhanika, 83-94.

[11] Rubinstein, R. Y. (1981). Simulation and the Monte Carlo Method. Wiley, New York.

[12] Sacks, J. (1958). Asymptotic distribution of stochastic approximation procedures. Ann. Math. Statist. 29, 373-405.

[13] Schruben, L. W. (1986). A frequency-domain approach to stochastic sensitivity analysis. Technical Report, Cornell University.

[14] Suri, R. (1983). Infinitesimal perturbation analysis of discrete event dynamic systems: A general theory. Proceedgings of the 22nd IEEE Conference on Decision and Control.

[15] Zazanis, M. A. and Suri, R. (1985). Comparison of perturbation analysis with conventional sensitivity estimates for regenerative stochastic systems. Technical Report, Harvard University.

BIOGRAPHY

Peter Glynn received his Ph.D. in Operations Research from Stanford University in 1982, and is currently an assistant professor at the University of Wisconsin-Madison. His research interests include queueing theory and computational probability. He is currently a member of ORSA, TIMS, IMS, and the Statistical Society of Canada.

Peter W. Glynn
Department of Industrial Engineering
and the Mathematics Research Center
University of Wisconsin-Madison
Madison, WI 53705, U.S.A.