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

STOCHASTIC APPROXIMATION FOR MONTE CARLO OPTIMIZATION

Peter W. Glynn
Department of Industrial Engineering
and Mathematics Research Center
University of Wisconsin-Madison
Madison, WI 53706, U.S.A.

ABSTRACT

In this paper, we introduce two convergent Monte Carlo algorithms for optimizing complex stochastic systems. The first algorithm, which is applicable to regenerative processes, operates by estimating finite differences. The second method is of Robbins-Monro type and is applicable to Markov chains. The algorithm is driven by derivative estimates obtained via a likelihood ratio argument.

1. INTRODUCTION

Our goal, in this paper, is to develop Monte Carlo algorithms that are capable of optimizing complex stochastic systems. By appropriately modifying classical stochastic approximation procedures, we are able to produce algorithms which are provably convergent.

Specifically, we consider the problem of optimizing the steady-state of a regenerative stochastic process with respect to a continuous decision parameter. We offer, in Section 2, a stochastic approximation algorithm for solving such problems, and specialize the procedure to Markov chains in Section 3. The proposed algorithm, while convergent under quite reasonable hypotheses, involves estimation of derivatives via finite differences. Since derivative estimation via finite differences is quite "noisy", one would anticipate that the algorithm converges rather slowly.

We therefore consider, in Section 4, a Robbins-Monro procedure which directly estimates the derivative without passing to a finite-difference approximation. In order to obtain such a derivative estimator, we find it necessary to impose additional Markov structure on the problem. We offer concluding remarks in Section 5.

2. STEADY-STATE OPTIMIZATION OF REGENERATIVE SYSTEMS

Let $X = \{X(t) : t \geq 0\}$ be a real-valued (possibly) delayed regenerative process with regeneration times $T(0) < T(1) < \dots$. Set $T(-1) = 0$ and let

$$Y_i = \int_{T(i-1)}^{T(i)} X(s)ds$$

$$\tilde{Y}_i = \int_{T(i-1)}^{T(i)} |X(s)|ds$$

$$\tau_i = T(i) - T(i-1) \quad .$$

It is well known (see SMITH (1955)) that if $E(\tilde{Y}_1 + \tau_1) < \infty$, then X is ergodic, in the sense that there exists a finite (deterministic) constant α such that

$$\frac{1}{t} \int_0^t X(s)ds \to \alpha \quad \text{a.s.}$$

as $t \to \infty$. The steady-state limit α can be expressed as a ratio of expectations:

$$\alpha = EY_1/E\tau_1 \quad .$$

Consequently, the estimation of the steady-state of a regenerative process can be viewed as a special case of the ratio estimation problem.

Motivated by this discussion, we shall now proceed to describe a stochastic optimization algorithm for the general ratio estimation problem. The goal, roughly speaking, will be to minimize α over some bounded open interval Λ; the probability distribution of the numerator and denominator random variables (r.v.'s) is dependent on the decision variable θ.

More specifically, assume $\Lambda = (-\pi/2, \pi/2)$. (Note that any optimization problem over (a,b) with $a < b$ can be transformed into an optimization problem over Λ by suitably re-parameterizing the decision variable θ.) The goal is to solve:

$$(2.1) \qquad \min_{\theta \in \Lambda} \alpha(\theta)$$

where $\alpha(\theta) = u(\theta)/\ell(\theta)$. We will assume that:

(2.2) $u(\theta) = EY(\theta)$, $\ell(\theta) = E\tau(\theta)$ when $(Y(\theta), \tau(\theta))$ has a distribution from which deviates can be generated, for $\theta \in \Lambda$

(2.3) $\ell(\theta) \neq 0$ for $\theta \in \Lambda$

(2.4) $u(\cdot)$, $\ell(\cdot)$ are twice continuously differentiable over Λ

(2.5) the derivatives $u^{(k)}(\cdot)$, $\ell^{(k)}(\cdot)$ $(k = 0,1,2,)$ are bounded over Λ

(2.6) $\alpha'(\theta^*) = 0$ has a unique root $\theta^* \in \Lambda$ and α is minimized there.

The assumptions (2.2) - (2.6) are, of course, of the kind that commonly appear in the analysis of deterministic optimization algorithms for solving nonlinear programming problems. Since we are proposing to use Monte Carlo methods to solve the optimization problem (2.1), it will be necessary to impose an additional "probabilistic" hypothesis on the distribution of $(Y(\theta), \tau(\theta))$.

$$(2.7) \qquad \sup_{\theta \in \Lambda} E(Y^2(\theta) + \tau^2(\theta)) < \infty \quad .$$

With (2.7) in hand, we can develop a stochastic approximation algorithm which converges to the minimizer θ^* associated with (2.1). We first recall that the standard stochastic approximation methods are formulated in terms of a decision variable λ which is uncontrained, in the sense that $\mathbf{R}$ is the corresponding parameter set. In order to transform $\mathbf{R}$ into Λ, we use the transformation:

$$(2.8) \qquad \theta = \arctan \lambda \quad ,$$

$\lambda \in \mathbf{R}$. (Relation (2.8) explains why it is convenient to choose $\Lambda = (-\pi/2, \pi/2)$.) Set $\underset{\sim}{\alpha}(\lambda) = \underset{\sim}{u}(\lambda)/\underset{\sim}{\ell}(\lambda)$ where

$$\underset{\sim}{u}(\lambda) = u(\arctan \lambda)$$
$$\underset{\sim}{\ell}(\lambda) = \ell(\arctan \lambda) \quad .$$

It is obvious that if one solves

$$(2.9) \qquad \min_{\lambda \in \mathbf{R}} \underset{\sim}{\alpha}(\lambda) \quad ,$$

then the minimizer λ^* to (2.9) corresponds (uniquely) to θ^*, and θ^* can be retrieved from (2.9) by setting $\theta^* = \arctan \lambda^*$. We shall therefore henceforth work with the optimization problem (2.9), in which the decision variable λ is unconstrained.

Because of the strict monotonicity of the arc tan function and assumption (2.6), it follows that λ^* is the unique root of

$$(2.10) \quad h(\lambda) = \underset{\sim}{u}'(\lambda)\underset{\sim}{\ell}(\lambda) - \underset{\sim}{u}(\lambda)\underset{\sim}{\ell}'(\lambda) \quad .$$

Thus, our goal can be re-expressed as: Find the unique root λ^* of $h(\cdot)$.

Since assumption (2.1) provides unbiased estimators $Y(\theta)$ $(\tau(\theta))$ only for function values $u(\theta)$ $(\ell(\theta))$ (and <u>not</u> for derivatives of u and ℓ), it is necessary to consider finite-difference approximations for the derivatives appearing in h. Then, for c small,

$$\underset{\sim}{u}'(\lambda) \approx c^{-1}[\underset{\sim}{u}(\lambda+c) - \underset{\sim}{u}(\lambda)]$$
$$\underset{\sim}{\ell}'(\lambda) \approx c^{-1}[\underset{\sim}{\ell}(\lambda+c) - \underset{\sim}{\ell}(\lambda)]$$

so that

$$(2.11) \quad h(\lambda) \approx c^{-1}[\underset{\sim}{u}(\lambda+c)\underset{\sim}{\ell}(\lambda) - \underset{\sim}{\ell}(\lambda+c)\underset{\sim}{u}(\lambda)] \quad .$$

We will drive our stochastic approximation algorithm with a Monte Carlo version of the right-hand side of (2.11). Let

$$\underset{\sim}{Y}(\lambda) = Y(\text{arc tan } \lambda)$$
$$\underset{\sim}{\tau}(\lambda) = \tau(\text{arc tan } \lambda) \quad .$$

Suppose that $(\underset{\sim}{Y}_1(\lambda+c), \underset{\sim}{\tau}_1(\lambda+c))$ and $(\underset{\sim}{Y}_1(\lambda), \underset{\sim}{\tau}_1(\lambda))$ are independent variates generated from the distributions of $(\underset{\sim}{Y}(\lambda+c), \underset{\sim}{\tau}(\lambda+c))$ and $(\underset{\sim}{Y}(\lambda), \underset{\sim}{\tau}(\lambda))$, respectively. Set

$$\eta = c^{-1}[\underset{\sim}{Y}_1(\lambda+c)\underset{\sim}{\tau}_1(\lambda) - \underset{\sim}{\tau}_1(\lambda+c)\underset{\sim}{Y}_1(\lambda)] \quad .$$

Because of the independence of the two pairs of variates, it is immediate that $E\eta$ is precisely the right-hand side of (2.11). Consequently, we shall use r.v.'s of the form η to drive our proposed Monte Carlo optimization algorithm.

Specifically, our algorithm shall take the form:

1.) Choose $\lambda_0 \in \mathbb{R}$.

2.) Given λ_n, generate (independently of the past) $(\underset{\sim}{Y}_{n+1}(\lambda_n + c_{n+1}), \underset{\sim}{\tau}_{n+1}(\lambda_n + c_{n+1}))$ and $(\underset{\sim}{Y}_{n+1}(\lambda_n), \underset{\sim}{\tau}_{n+1}(\lambda_n))$ independently from the distributions of $(\underset{\sim}{Y}(\lambda_n + c_{n+1}), \underset{\sim}{\tau}(\lambda_n + c_{n+1}))$ and $(\underset{\sim}{Y}(\lambda_n), \underset{\sim}{\tau}(\lambda_n))$, respectively, where $\{c_n : n \geq 1\}$ is a sequence of positive constants decreasing to zero.

3.) Set

$$\eta_{n+1} = c_{n+1}^{-1}[\underset{\sim}{Y}_{n+1}(\lambda_n + c_{n+1})\underset{\sim}{\tau}_{n+1}(\lambda_n) - \underset{\sim}{Y}_{n+1}(\lambda_n)\underset{\sim}{\tau}_{n+1}(\lambda_n + c_{n+1})] \quad .$$

4.) Then, let

$$\lambda_{n+1} = \lambda_n - a_{n+1}\eta_{n+1} \tag{2.12}$$

where $\{a_n : n \geq 1\}$ is a sequence of positive deterministic constants satisfying

$$\sum_{n=1}^{\infty} a_n = \infty, \quad \sum_{n=1}^{\infty} a_n^2 < \infty, \quad \sum_{n=1}^{\infty} a_n c_n < \infty,$$
$$\sum_{n=1}^{\infty} a_n^2/c_n^2 < \infty \quad .$$

5.) Go to 2.).

Note that, as expected, the iteration (2.12) shows that λ_{n+1} will tend to be smaller (larger) than the "current" value λ_n if $h(\lambda_n)$ is positive (negative). This is clearly the behavior required, in order that λ_n converge to a point λ^* at which $h(\lambda^*+) > 0$, $h(\lambda^*-) < 0$. These latter conditions on $h(\lambda^*)$ essentially dictate that λ^* is a local minimizer of $\underset{\sim}{\alpha}$. We can now state our convergence result.

(2.14). THEOREM. Under assumptions (2.2) - (2.7), $\lambda_n \to \lambda^*$ a.s. as $n \to \infty$, where λ^* is the solution of (2.9).

For a proof, see Appendix 1. Reasonable choices for $\{a_n\}$, $\{c_n\}$ are $a_n = a/n$ $(a>0)$, $c_n = c/n^{1/4}$ $(c>0)$.

3. FINITE DIFFERENCE ALGORITHMS FOR MARKOV CHAINS

Our goal, in this section, is to apply the finite difference algorithm of Section 2 to the steady-state optimization of Markov chains.

For $\theta \in \Lambda = (-\pi/2, \pi/2)$, let $P(\theta)$ be an irreducible $d \times d$ stochastic matrix. Since $P(\theta)$ is irreducible and finite, it follows that $P(\theta)$ has a unique stationary distribution $\pi(\theta)$ satisfying

$$\pi(\theta)^t P(\theta) = \pi(\theta)^t \quad . \tag{3.1}$$

(We follow the convention that all d-vectors are written as column vectors.) Suppose that $f(\theta)$ is a performance function in which $f(\theta,i)$ denotes the cost incurred by the chain in state i, over one unit of time, when driven by $P(\theta)$. If $X = \{X_n : n \geq 0\}$ is a Markov chain evolving under $P(\theta)$, then

$$\frac{1}{n}\sum_{k=0}^{n-1} f(\theta,X_k) \to \alpha(\theta) \qquad P_\theta \text{ a.s.}$$

(P_θ is the probability on the path-space of X induced by $P(\theta)$) where $\alpha(\theta)$ is the deterministic constant

$$(3.2) \qquad \alpha(\theta) = \sum_{i\in S} \pi(\theta,i)f(\theta,i) \quad .$$

(S is the state space of X.) The goal here is to obtain a version of the algorithm stated in Section 2, that is capable of solving:

$$(3.3) \qquad \min_{\theta\in\Lambda} \alpha(\theta) \quad .$$

Before proceeding further, it should be observed that deterministic non-Monte Carlo algorithms are available for solving (3.3). One could, for example, apply a mathematical programming algorithm based on function evaluations of $\alpha(\theta)$. Each function evaluation would involve solving (3.1) subject to $\pi(\theta)^t e = 1$ (e is the vector consisting entirely of 1's), followed by computation of the inner product (3.2). Because of the intrinsic slow convergence of Monte Carlo algorithms relative to standard nonlinear programming procedures, one would expect that a deterministic procedure of the type described above would be preferable to a Monte Carlo algorithm. This is a reasonable expectation when function evaluations aren't too expensive. However, for large d, finding $\alpha(\theta)$ is computationally intensive and Monte Carlo procedures become attractive.

The assumptions that we shall require on $P(\theta)$ and $f(\theta)$ are as follows:

(3.4) $P(\theta)$ is irreducible for $\theta \in \Lambda$.

(3.5) $P(\cdot)$, $f(\cdot)$ are twice continuously differentiable on $[-\pi/2, \pi/2]$.

(3.6) $\alpha'(\theta^*) = 0$ has a unique root $\theta^* \in \Lambda$ and α is minimized there.

(3.7) There exists a state $i \in S$ such that $P(\theta,k,i)$ is positive for $\theta \in [-\pi/2, \pi/2]$, $k \in S$.

In order to apply the regenerative methodology of Section 2, it is necessary to observe that for all $\theta \in \Lambda$, X is regenerative under P_θ with regeneration times consisting of the sequence of times $\{T(n) : n \geq 0\}$ on which i is hit. Consequently,

$$\alpha(\theta) = EY(\theta)/E\tau(\theta)$$

where

$$Y(\theta) = \sum_{k=0}^{T(\theta)-1} f(\theta,X_k(\theta))$$

$$\tau(\theta) = T(\theta)$$

$$T(\theta) = \inf\{k \geq 1 : X_k(\theta) = i\}$$

and $X(\theta)$ is a Markov chain on S evolving under $P(\theta)$, with $X_0(\theta) = i$. We can now apply the algorithm of Section 2 to solving (3.3).

ALGORITHM A:

1. Choose $\theta_0 \in \Lambda$.
2. Given θ_n, generate a trajectory of $X(\theta_n)$, with initial condition $X_0(\theta_n) = i$, until the first hitting time of i. Set

$$Y_{n+1} = \sum_{k=0}^{T(\theta_n)-1} f(\theta_n,X_k(\theta_n))$$

$$\tau_{n+1} = T(\theta_n)$$

$$T(\theta_n) = \inf\{k \geq 1 : X_k(\theta_n) = i\}.$$

3. Given θ_n, generate a trajectory of $X(\theta_n')$, with initial condition $X_0(\theta_n') = i$, until the first hitting time of i, where $\theta_n' = \arctan(\tan(\theta_n) + n^{1/4})$.

Set

$$Y_{n+1}' = \sum_{k=0}^{T(\theta_n')-1} f(\theta_n',X_k(\theta_n'))$$

$$\tau_{n+1}' = T(\theta_n')$$

$$T(\theta_n') = \inf\{k \geq 1 : X_k(\theta_n') = i\}.$$

4. Compute

$$\eta_{n+1} = n^{1/4}[Y'_{n+1}\tau_{n+1} - Y_{n+1}\tau'_{n+1}] \quad .$$

5. Compute

$$\theta_{n+1} = \arctan[\tan(\theta_n) - (1/n)\eta_{n+1}] \quad .$$

6. Go to 2.

Theorem 2.14 can be applied to obtain a convergence result for this algorithm.

(3.8) THEOREM. Under (3.4) - (3.7), the θ_n's produced by Algorithm A converge a.s. to θ^*, where θ^* solves (3.3).

For a proof, see Appendix 2. Note that assumptions (3.4), (3.5), and (3.7) are easy to check. Hypothesis (3.6) is more implicit, but rather typical of assumptions commonly imposed in the analysis of nonlinear programming procedures.

4. A ROBBINS-MONRO ALGORITHM FOR OPTIMIZATION OF MARKOV CHAINS

Our goal is to solve (3.3) by applying a stochastic approximation algorithm which replaces the finite difference approximation appearing in (2.10) by an unbiased Monte Carlo estimator of the derivative. The expectation is that the convergence rate of a derivative-driven algorithm should be considerably faster than that associated with a finite difference estimator.

In order to accomplish this task, we shall require unbiased estimators for the terms $\underset{\sim}{u}'(\lambda)$, $\underset{\sim}{\ell}'(\lambda)$ appearing in (2.10). Recall that in the Markov chain context of interest here,

$$u(\theta) = EY(\theta)$$

$$\ell(\theta) = E\tau(\theta)$$

where $Y(\theta)$, $\tau(\theta)$ are r.v.'s formed by simulating $X(\theta)$ over one regenerative cycle. We shall apply likelihood ratio techniques to obtain a suitable estimator of the derivatives. This will require recasting the way in which we view the Markov chain's dependence on θ.

Specifically, we have so far (implicitly) viewed the dependence in terms of constructing a different process $X(\theta)$ for each $\theta \in \Lambda$. Another way to approach the problem is to view the Markov chain in terms of a single process X whose distribution on path-space depends on θ. To make this rigorous, let $\Omega = S \times S \times \ldots$, and let F be the associated product σ-field. Let $X = \{X_n : n \geq 0\}$ be the co-ordinate process defined by $X_n(\omega) = \omega_n$, where $\omega = (\omega_0, \omega_1, \ldots)$ is a typical element of Ω. Note that for each $\theta \in \Lambda$, the transition matrix $P(\theta)$ induces a probability P_θ on the path-space (Ω, F) via:

$$\text{(4.1)} \quad P_\theta\{X_0 = i_0, X_1 = i_1, \ldots, X_n = i_n\} = \delta_{i,i_0} P(\theta, i_0, i_1) \cdots P(\theta, i_{n-1}, i_n) \quad .$$

In order to apply the likelihood ratio idea, it is convenient to assume that:

(4.2) $\Gamma(\theta) = \{(j,k) : P(\theta,j,k) > 0\}$ is independent of θ, for $\theta \in \Lambda$.

Assumption (4.2) allows one to multiply and divide by $P(\theta_0,j,k)$ in (4.1) to obtain

$$\text{(4.3)} \quad P_\theta\{X_0 = i_0, \ldots, X_n = i_n\} = E_{\theta_0}\{L(n,\theta,\theta_0); X_0 = i_0, \ldots, X_n = i_n\}$$

where

$$L(n,\theta,\theta_0) = \prod_{k=0}^{n-1} \frac{P(\theta, X_k, X_{k+1})}{P(\theta_0, X_k, X_{k+1})}$$

is the "likelihood ratio" associated with X. An easy argument, based on the observation that $T = \inf\{n \geq 1 : X_n = i\}$ is a stopping time, shows that

$$\text{(4.4)} \quad u(\theta) = E_{\theta_0} Y(f(\theta)) L(T,\theta,\theta_0)$$

where

$$Y(g(\theta)) = \sum_{k=0}^{T-1} g(\theta, X_k) \quad .$$

If one formally differentiates (4.4) by interchanging the derivative and expectation, one obtains

$$(4.5) \qquad u'(\theta_0) = E_{\theta_0} Y(f'(\theta_0)) + E_{\theta_0} Y(f(\theta_0)) L'(T, \theta_0, \theta_0)$$

where

$$L'(T, \theta_0, \theta_0) = \sum_{k=0}^{T-1} \frac{P'(\theta_0, X_k, X_{k+1})}{P(\theta_0, X_k, X_{k+1})} \quad .$$

The right-hand side of (4.5) can be used to obtain Monte Carlo estimates of $u'(\theta)$; a similar estimator exists for $\ell(\theta)$ (set $f(\theta) \equiv 1$ in (4.5)). To justify the interchange we require that:

(4.6) $P(\cdot)$, $f(\cdot)$ are continuously differentiable over $[-\pi/2, \pi/2]$.

(4.7) $P(\theta)$ is irreducible, for $\theta \in \Lambda$.

The hypotheses (4.2), (4.6), (4.7) are easily checked, in most applications. Unfortunately, in order to obtain an (easy) convergence proof for our stochastic approximation algorithm, we shall also need the more implicitly stated:

(4.8) $\alpha'(\theta^*) = 0$ has a unique root $\theta^* \in \Lambda$ and α is minimized there.

We can now state our stochastic approximation derivative-based algorithm.

ALGORITHM B:

1. Choose $\theta_0 \in \Lambda$, $i \in S$.
2. Given θ_n, generate a trajectory of X with transition matrix $P(\theta_n)$ and initial condition $X_0 = 1$, until the first hitting time of i. Set

$$Y_{n+1} = \sum_{k=0}^{T_{n+1}-1} f(\theta_n, X_k)$$

$$T_{n+1} = \inf\{k > 1 : X_k = i\}$$

$$D_{n+1} = \sum_{k=0}^{T_{n+1}-1} f'(\theta_n, X_k) + Y_{n+1} \cdot \sum_{k=0}^{T_{n+1}-1} \frac{P'(\theta_n, X_k, X_{k+1})}{P(\theta_n, X_k, X_{k+1})}$$

$$E_{n+1} = T_{n+1} \cdot \sum_{k=0}^{T_{n+1}-1} \frac{P'(\theta_n, X_k, X_{k+1})}{P(\theta_n, X_k, X_{k+1})} \quad .$$

3. Repeat Step 2 with an independent trajectory, obtaining Y'_{n+1}, T'_{n+1}, D'_{n+1}, E'_{n+1}.
4. Compute

$$\eta_{n+1} = (2 + 2\tan^2\theta_n)^{-1}[D_{n+1}T'_{n+1} + D'_{n+1}T_{n+1} - E_{n+1}Y'_{n+1} - E'_{n+1}Y_{n+1}] \quad .$$

5. Compute

$$\theta_{n+1} = \arctan\,[\tan(\theta_n) - (1/n)\eta_{n+1}] \quad .$$

6. Go to 2.

A convergence result is also available for this algorithm.

(4.9) THEOREM. Assume (4.2) and (4.6) - (4.8). Then, the sequence θ_n produced by Algorithm B converges a.s. to θ^*, where θ^* solves (3.3).

See Appendix 3 for a proof.

Note that because of the independence of the observations generated in Step 2 and 3, $E\eta_{n+1} = h(\theta_n)$ where h is the function defined by (2.10). Because the stochastic approximation procedure described here attempts to find a zero of h using unbiased estimators of $h(\theta_n)$, this algorithm is of Robbins-Monro type.

5. CONCLUSION

We have discussed two optimization algorithms in this paper. The first algorithm, which requires only regenerative structure, involves finite-differences and is expected to be rather slowly convergent. The second algorithm, by exploiting Markov structure, uses derivative estimates, and is of Robbins-Monro type. More work is needed in this area, particularly empirical investigations.

ACKNOWLEDGEMENTS

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

APPENDIX 1: PROOF OF THEOREM 2.14

We will apply a convergence theorem due to Metivier and Priouret (1984). Observe that

$$\lambda_{n+1} = \lambda_n - a_{n+1}h(\lambda_n) + a_{n+1}\beta_{n+1} - a_{n+1}\varepsilon_{n+1}$$

where

$$\beta_{n+1} = h(\lambda_n) - c_{n+1}^{-1}[\underset{\sim}{u}(\lambda_n+c_{n+1})\underset{\sim}{\ell}(\lambda_n) - \underset{\sim}{u}(\lambda_n)\underset{\sim}{\ell}(\lambda_n+c_{n+1})]$$

$$\varepsilon_{n+1} = \eta_{n+1} - c_{n+1}^{-1}[\underset{\sim}{u}(\lambda_n+c_{n+1})\underset{\sim}{\ell}(\lambda_n) - \underset{\sim}{u}(\lambda_n)\underset{\sim}{\ell}(\lambda_n+c_{n+1})] .$$

To apply Theorem C of [2], we observe that the monotonicity of the arc tan function and (2.6) imply that

$$\frac{dx}{dt} = -h(x(t))$$

has the unique equilibrium point λ^* and the domain of attraction of λ^* is $\mathbb{R}$. We now need to verify that $\beta_n \to 0$ a.s. as $n \to \infty$. By using Taylor's expansion to two terms, we find that

$$\beta_{n+1} = -\frac{c_{n+1}}{2}[\underset{\sim}{u}''(\xi_{n+1})\underset{\sim}{\ell}(\lambda_n) - \underset{\sim}{u}(\lambda_n)\ell''(\xi_n)]$$

where ξ_{n+1} lies between λ_n and $\lambda_n + c_{n+1}$. Now,

$$\underset{\sim}{u}''(x) = \frac{u''(\text{arc tan } x) - u'(\text{arc tan } x)(2x)}{(1+x^2)^2} .$$

Since $u''(\text{arc tan } x)$, $u'(\text{arc tan } x)$ are uniformly bounded in x (see (2.5)), it follows that $\underset{\sim}{u}''(\cdot)$ is a bounded function. A similar argument shows that $\underset{\sim}{\ell}''(\cdot)$ is bounded. The boundedness of $\underset{\sim}{u}$ and $\underset{\sim}{\ell}$ (see (2.5)) therefore implies the existence of a deterministic constant β such that

$$|\beta_{n+1}| \le \beta c_{n+1} .$$

Thus, $\beta_n \to 0$ a.s. as $n \to \infty$.

To verify (H2) of Theorem C of [2], observe that the sequence $\{\varepsilon_n : n \ge 1\}$ constitute martingale differences, and apply Doob's inequality. The proof is complete if we can show that $\sup\{|\lambda_n| : n \ge 0\} < \infty$ a.s. The quasimartingale argument given on p. 143 of [2] is valid here also because of (2.13) and

$$E\eta_{n+1}^2 \le \frac{4}{c_{n+1}^2}\,[\sup_{\theta\in\Lambda} E(Y^2(\theta) + \tau^2(\theta))]^2 .$$

The conditions of Theorem C being in force, we may conclude that $\lambda_n \to \lambda^*$ as $n \to \infty$.

APPENDIX 2: PROOF OF THEOREM 3.8

We just need to check that the hypotheses (3.4) - (3.7) translate into (2.2) - (2.7), in this problem setting. Of course, (2.2) is immediate. For (2.3), observe that $T(\theta) \ge 1$ a.s. for all θ, so $\ell(\theta) \ge 1$. We refer to Appendix 3 for a proof of (2.4) - (2.5). (The proof given there for the first derivative easily extends to the second derivative.) Assumptions (2.6) and (3.6) are identical, leaving (2.7) to be validated. Let

$$\delta = \inf\{P(\theta,k,i) : \theta \in \Lambda,\ k \in S\} .$$

Since $P(\cdot,k,i)$ is continuous over $[-\pi/2,\pi/2]$ (see (3.5)), it follows from the positivity condition (3.7) and the finiteness of S that $\delta > 0$. Since

$$P\{X_{n+1}(\theta) \neq i \mid X_n(\theta)\} = 1 - P(\theta,X_n(\theta),i) \le 1 - \delta ,$$

it is evident that

$$P\{T(\theta) > n\} = P\{X_1(\theta) \neq i,\ldots,X_n(\theta) \neq i\} \le (1-\delta)^n .$$

Clearly, $|Y(\theta)| \le \|f(\theta)\|T(\theta)$ ($\|f(\theta)\| = \max\{|f(\theta,i)| : i \in S\}$), so

$$E(Y^2(\theta) + \tau^2(\theta)) \le (\|f(\theta)\|^2 + 1)ET^2(\theta)$$

$$\le 2(\|f(\theta)\|^2 + 1)\sum_{n=1}^{\infty} n\, P\{T(\theta) > n\}$$

$$\le 2(\|f(\theta)\|^2 + 1)\sum_{n=1}^{\infty} n(1-\delta)^{n-1}$$

(see p. 44 of Chung (1974)). Since $\|f(\theta)\| \le \sup\{|f(\theta,k)| : \theta \in \Lambda,\ k \in S\} < \infty$ (use (3.5) and the finiteness of S), (2.7) follows easily.

APPENDIX 3: PROOF OF THEOREM 4.9

We first need to verify that the interchange of derivative and expectation required to obtain (4.5) is valid. To do this, we start by showing that T has a moment generating function which converges in a neighborhood of zero.

First, let $\Gamma = \Gamma(\theta)$ and observe that by (4.6)

$$\inf\{P(\theta,j,k) : \theta \in \Lambda\} = R(j,k) > 0$$

precisely when $(j,k) \in \Gamma$. Hence, by (4.7),

$$P(\theta) \ge R$$

for all $\theta \in \Lambda$, where R is irreducible. (Irreducibility is a notion which makes sense for any non-negative matrix.) Since R is irreducible, it follows that

$$(R + R^2 + \ldots + R^d)$$

is a strictly positive matrix. Thus, it is evident that

$$\inf\{\sum_{\ell=1}^{d} P^{\ell}(\theta,j,k) : \theta \in \Lambda\} > 0$$

for all pairs $(j,k) \in S \times S$. For any $i \in S$, the finiteness of S implies that

$$\inf\{\sum_{\ell=1}^{d} P^{\ell}(\theta,k,i) : \theta \in \Lambda,\ k \in S\} = \delta > 0$$

and hence, for $\ell \le d$,

$$P_\theta\{T \le d \mid X_0\} \ge P_\theta\{X_\ell = i \mid X_0\} = P^{\ell}(\theta,X_0,i)$$

so

$$P_\theta\{T \le d \mid X_0\} \ge \max\{P^{\ell}(\theta,k,i) : k \in S\} \ge \delta/d .$$

It follows that

$$P_\theta\{T > kd\} \le (1 - \delta/d)^k ,$$

and hence T has a moment generating function which converges for arguments t such that $t < -\ell n(1 - \delta/d)$.

To justify the interchange, we will show that the difference quotients

$$h^{-1}[Y(f(\theta_0+h))L(T,\theta_0+h,\theta_0) - Y(f(\theta_0))]$$

are dominated by an integrable r.v. By the mean value theorem, the difference quotient equals

$$Y(f'(\eta))L(T,\eta,\theta_0) + Y(f(\eta))L'(T,\eta,\theta_0) \quad .$$

Clearly,

$$|Y(f'(\eta))| < \|f'\| \cdot T$$
$$|Y(f(\eta))| < \|f\| \cdot T$$

where $\|f'\| = \sup\{|f'(\theta,k)| : \theta \in \Lambda,\ k \in S\}$, $\|f\| = \sup\{|f(\theta,k)| : \theta \in \Lambda,\ k \in S\}$ are finite, by (4.6). Also,

$$L(T,\eta,\theta_0) < \varphi(h)^T$$

where

$$\varphi(h) = \max\{\frac{P(\theta,i,j)}{P(\theta_0,i,j)} : |\theta - \theta_0| < h, \quad (i,j) \in \Gamma\}$$

and

$$L'(T,\eta,\theta_0) < \|p'\| \cdot T \cdot \varphi(h)^T \ ;$$

$$\|p'\| = \sup\{P'(\theta,i,j)/P(\theta_0,i,j) : \theta \in \Lambda \ , \quad (i,j) \in \Gamma\} \ .$$

By (4.6), $\|p'\| < \infty$ and $\varphi(h) \to 1$ as $h \to 0$. So, our difference quotients are dominated by

$$(\|f'\| \cdot T + \|f\| \cdot \|p'\| \cdot T^2)\varphi(h)^T \quad .$$

Since T has a convergent moment generating function, it is evident that the above r.v. is integrable for h sufficiently small (use the fact that $\varphi(h) \to 1$), proving the required domination. Hence, $u(\cdot)$ is differentiable, with the derivative represented by the right-hand side of (4.5). A similar argument shows $\ell(\cdot)$ is differentiable (set $f(\theta) \equiv 1$).

We now wish to apply the results of Section II.F of [2], in order to obtain convergence of the algorithm itself. Assumptions (RM1) - (RM4) are automatic. To finish the proof, we need to verify condition (RM5) of [2] and show $h(\cdot)$ is continuous. To obtain (RM5), note that

$$|Y_{n+1}| < \|f\| \cdot T_{n+1}$$

$$|D_{n+1}| < \|f'\| \cdot T_{n+1} + \|f\|\cdot\|p'\| \cdot T_{n+1}^2$$

$$|E_{n+1}| < \|p'\| \cdot T_{n+1}^2$$

so

$$\eta_{n+1}^2 < 32(1 + \|f\|^2 + \|f'\|^2 + \|p'\|^2)^2 \cdot T_n^4(T'_{n+1})^4$$

where T'_{n+1} is an independent copy of T_{n+1}. Thus, the variance of η is bounded above if

$$\sup_\theta E_\theta T^4 < \infty$$

which follows since the tail of T is uniformly bounded in θ:

$$E_\theta T^4 < 4 \sum_{n=1}^{\infty} n^3 d^3 (1 - \delta/d)^{n-1} \ ,$$

proving (RM5).

We finish by observing that the continuity of h follows from showing u, u', ℓ, ℓ' are continuous. But an easy dominated convergence argument, using the bounds on Y, T, D, E already shown, provides the required proof, completing the argument.

REFERENCES

Chung, K. L. (1974). A Course in Probability Theory. Academic Press, New York.

Metivier, M. and Priouret, P. (1984). Applications of a Kushner and Clark lemma to general classes of stochastic algorithms. IEEE Trans. Inform. Theory 30, 140-151.

Smith, W. L. (1955). Regenerative stochastic processes. Proc. Roy. Soc. A232, 6-31.

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 Mathematics Research Center
University of Wisconsin-Madison
Madison, WI 53705, U.S.A.