ITERATIVE SIMULATION USING SINGLE AND MULTIPLE SEQUENCES

499

replications before listing several points and references that I hope are new to some readers.

I enjoyed both papers, but my general preference remains unchanged: Use a single long replication, except in special cases such as parallel computing or when stratified or antithetic initial states happen to be easy to determine. This preference for a single replication is due to its robustness to analyst lack of sophistication or time. Fifteen years ago substantial background, insight, and effort were required for simulation and for statistics practitioners to analyze complex problems. Commercial software has blossomed in both fields, allowing relatively naive practitioners to expect something good to happen when they give their problem to the computer. Similarly, one day we will expect software to evaluate posterior distributions with little practitioner insight. The single long replication makes negligible the initial bias, thereby alleviating the difficult initial-data-deletion problem.

Glynn and Heidelberger (1992) and Kelton (1989) are recent additions to the extensive literature that discusses initial deletion of warm-up data and choice of initial states.

Glynn (1987), Whitt (1990) and Damerdji (1991) discuss the choice of number of replications, the extreme cases being a single long run and many short runs.

Smith (1984) discusses Monte Carlo sampling from doubly stochastic Markov chains. The motivation is the need to identify nonredundant constraints in mathematical programming. The methods can be used to sample from a density by sampling uniformly within the region defined by the density and the zero plane. The Hit-and-Run sampler (Belsile, Romeijn and Smith 1992) is a generalization to nonuniform distributions.

Since essentially all point estimators are asymptotically normal, sampling error is well summarized by point-estimator standard error. The method of nonoverlapping adjacent batch means (NBM) is extended to overlapping batch means (OBM) in Maketon and Schmeiser (1984). OBM are highly dependent, which is acceptable since batches are sufficiently large not when the batch means are essentially independent but (loosely) when each batch subsumes the autocorrelation structure.

Except for end effects, OBM is the Bartlett-window spectral estimator with lag-window length equal to the batch size. Therefore, OBM has the same bias but only two-thirds the variance of NBM. Both NBM and OBM estimator are easily computed in $O(n)$ time; therefore OBM dominates NBM for Markov chain sampling.

For NBM, OBM and some other estimators based on batching, the mse-optimal batch size is asymptotically

$$m^* = \left[ \frac{2n}{c_0 (c_0/\gamma_0)^2} \right]^{1/3},$$

where $c_0$ and $c_0$ are the estimator's bias and variance constants, respectively, and $\gamma_1/\gamma_0$ is the center of gravity of the absolute values of the autocorrelation lags. For OBM $c_0 = 1$ and $c_0 = 4/3$. Geyer’s Theorem 3.1 helps to estimate the autocorrelation center of gravity, which is problem dependent, but the goal is to estimate optimal batch size without estimating individual autocorrelations.

An advantage of batch-means methods is that they extend directly to point estimators that are not means. Schmeiser, Avramidis and Hashem (1990) discuss sufficient assumptions and provide a code for overlapping batch variances and overlapping batch quantiles. Nelson (1989) quantifies the additional number of batches needed when estimating optimal weights for control variates.

For random-number and random-variate generation, see Fishman and Moore (1986) and Devroye (1986), respectively.

Glasserman (1991) discusses single-replication methods for estimating derivatives of performance measures with respect to system design parameters. Could similar methods be used to estimate the change, for example, in the posterior mean caused by a unit change in the prior mean?

A variety of other simulation-experiment issues are discussed in Schmeiser (1990).

Comment

Luke Tierney

Both papers make some interesting contributions to the discussion of issues related to Markov chain Monte Carlo. Geyer’s variance estimates that take advantage of the Markov chain structure appear to be particularly promising and worthy of further investigation. As these methods require a reversible chain, they are not directly applicable to the fixed scan Gibbs sampler. But several simple devices are available for making Gibbs samplers reversible, including random scans.

Luke Tierney is Professor, School of Statistics, University of Minnesota, Vincent Hall, Minneapolis, Minnesota 55455.