A CROSS-SPECTRAL METHOD FOR SENSITIVITY ANALYSIS
OF COMPUTER SIMULATION MODELS

A. Feuerverger, D.L. McLeish and R. Rubinstein

Presented by D.R. Brillinger, F.R.S.C.

ABSTRACT. Cross-spectral analysis of system performance variables with
correlation function sequences leads to a practical solution of the problem of sensi-
tivity analysis for computer simulation models.

1. INTRODUCTION. Consider a computer simulation model driven by an in-
put sequence \( X_t \) and resulting in an output sequence \( Y_t \) where
\( t = 0, \pm 1, \pm 2, \ldots \). The input sequence \( X_t \) is taken as independent identi-
cally distributed from density \( f_\nu(x) \) where \( \nu \) is a multidimensional real
parameter, and the output sequence \( Y_t \) will normally settle (as \( t \to \infty \)) into
steady state and become a stationary and ergodic process. One or more sample
performance measures of the form \( L_\nu = L_\nu(Y_t) \) are evaluated and we are in-
terested not only in the steady state mean \( \bar{l}(\nu) = \lim_{t \to \infty} E_\nu L_\nu \) but also in the
sensitivities (gradient, Hessian, etc.) \( \nabla_\nu \bar{l}(\nu), \nabla_\nu^2 \bar{l}(\nu) \). Examples of relevant
stochastic systems are queuing and reliability networks. In the first case \( L_\nu \)
might be the sojourn time of the \( t \)-th customer, and \( f_\nu(x) \) the multivariate
density of interarrival times, service times and routing probabilities. In the
second case \( L_\nu \) might be the life of a reliability system while \( f_\nu(x) \) describes
the component lifetimes. In such systems \( \bar{l}(\nu) \) is generally not analytically
tractable so that we have to resort to Monte Carlo simulation. Normally the
system cannot have knowledge of the future and we write
A. Feuerverger, D.L. McLeish and R. Rubinstein

\[ Y_i = Y_i(X_i, X_{i-1}, \ldots) \text{ and } L_i = L_i(Y_i) = L_i(X_i, X_{i-1}, \ldots). \]

In typical applications, system operation is started from some initial state and left to run until stationarity is attained. Thereafter \( T \) consecutive observations are taken and we denote these as \((X_1, L_1), (X_2, L_2), \ldots, (X_T, L_T)\). In this simulation the value of \( \nu \) is set at \( \nu_0 \) and due to the complexity of the system, it is costly to repeat the simulation at other values of \( \nu \). The value of \( I(\nu_0) \) may be estimated by $\frac{1}{T} \sum_{t=4}^{T} L_t$ whose variance, under general conditions, is \( O(T^{-1}) \).

The purpose of this note is to present an effective method by which the sensitivities may be estimated simultaneously from the same simulation run. Some relevant references are Rubinstein (1986), Ho and Cao (1983). One contribution of our new method lies in the substantial reduction in the asymptotic order of the variance achieved relative to the score function method (Rubinstein, 1986), namely from \( O(T) \) to \( O(B_T^{-1} T^{-1}) \) for \( B_T \to 0 \) such that \( B_T T \to \infty \).

2. MAIN RESULT. For simplicity we take \( \nu \) here to be univariate and assume the process \( L_t \) is stationary. Application of the result to vector valued parameters \( \nu \) and to a vector of performance measures \( L_t \) requires only considering the sequences \( L_t \) and \( S_t \) appearing in the theorem to be jointly stationary vector valued time series. Hereafter the parameter \( \nu \) will be assumed to be set equal to its value in the simulation \( \nu_0 \) wherever it appears. Our key result requires a mixing type condition consistent with the physical requirement that the simulation system settles eventually into a steady state suitable for statistical analysis. Specifically, let \( F_{\nu}^t \) be the \( \sigma \)-field

\[ F_{\nu}^t = \sigma \{ X_{M}, X_{M+1}, \ldots, X_t \}. \]

It follows from the martingale convergence theorem (e.g. Doob, 1953, p. 331, Theorem 4.3) that the approximation to the function \( L_t \) based on a finite data set \( X_{M}, X_{M+1}, \ldots, X_t \), say, approxi-
mates $L_t$ arbitrarily closely as $M \to \infty$. More precisely, if $L_t$ is measurable with respect to $F^{t}_\infty$ and is square integrable, then

$$E_\nu[L_t | F_{-M}^t] \to L_t \text{ as } M \to \infty$$

(2.1)

where the convergence in (2.1) holds both with probability 1 and in expectation since the martingale on the left hand side of (2.1) is uniformly integrable since it has a bounded sequence of second moments. Now if we are to be able to approximate the sensitivity $\frac{\partial}{\partial \nu} l(\nu)$ also using only a finite data set, then clearly it is necessary that the convergence in (2.1) occur for the expectation of the derivative with respect to $\nu$ as well, i.e. that

$$0 = E_\nu \frac{\partial}{\partial \nu} L_t = \lim_{M \to \infty} E_\nu \left\{ \frac{\partial}{\partial \nu} E_\nu [L_t | F_{-M}^t] \right\}.$$  

(2.2)

Our main result is the following

**THEOREM.** Let $(X_\nu, L_t)$ be the stationary stochastic system described above. Let $S_t = \frac{\partial \log f_\nu(X_t)}{\partial \nu}$ and assume that $S_t$ and $L_t$ are square integrable and that the covariances $\text{cov}_\nu(L_u, S_v)$ are absolutely summable in $t$. Assume that (2.2) holds and also that the derivative may be passed through the integral in

$$\frac{\partial}{\partial \nu} \int E_\nu[L_t | F_{-M}^t] \prod_{i=-M}^{1} f_\nu(x_i) \, dx$$

for each $M$. Then

$$\frac{\partial}{\partial \nu} l(\nu) = f_{L_S}(0)$$

(2.3)

where $f_{L_S}(\lambda)$ is the cross-spectral density function of the stationary sequence $(L_u, S_t)$. It follows that for any sequence $B_T \to 0$ such that $B_T T \to \infty$, an estimator will exist having bias $O(B_T^2)$ and variance $O(B_T^{-1} T^{-1})$. 
PROOF. From (2.2)

\[
\frac{\partial}{\partial \nu} l(\nu) = \lim_{M \to \infty} \frac{\partial}{\partial \nu} E_{\nu}(L_0 | F^1_M) \tag{2.4}
\]

\[
= \lim_{M \to \infty} \int \frac{\partial}{\partial \nu} E_{\nu}(L_0 | F^1_M) \prod_{-M}^{1} f_\nu(x_i) \, dx + \int E_{\nu}(L_0 | F^1_M) \frac{\partial}{\partial \nu} \prod_{-M}^{1} f_\nu(x_i) \, dx \tag{2.5}
\]

\[
= \lim_{M \to \infty} E_{\nu} \left( \frac{\partial}{\partial \nu} E_{\nu}(L_0 | F^1_M) \right) + E_{\nu}(L_0 \sum_{-M}^{1} S_i) \tag{2.6}
\]

\[
= \sum_{j=0}^{\infty} \text{cov}_{\nu}(L_0, S_j) \tag{2.7}
\]

\[
= f_{L_0}(0). \tag{2.8}
\]

The proof shows that the limit in (2.2) must exist under the other conditions of the theorem; (2.2) is only required to ensure that this limit is 0. Concerning estimation of \( f_{L_0}(0) \), see Brillinger (1975), Jenkins and Watts (1969). In particular (Brillinger, 1975, chapter 7) for any sequence \( B_T \to 0 \) such that \( B_T T \to \infty \) an estimator will generally exist having bias \( O(B_T) \) and variance \( O(B_T^{-1}T^{-1}) \); for symmetric weight functions the bias will be \( O(B_T^2) \). Variance reducing techniques which take into account the onesidedness of the \( f_{L_0}(\lambda) \) Fourier series (e.g. Bhansali and Karavellas, 1983) and as in Heidelberger and Welch (1981) may also be applied. Further terms in the Taylor expansion of \( l(\nu) \) may be obtained by means of cumulant spectra (Brillinger and Rosenblatt, 1967), e.g.

\[
\frac{\partial^2}{\partial \nu^2} l(\nu) = f_{L_0,S}(0,0) \quad \text{where} \quad f_{L_0,S}(\lambda_1, \lambda_2) \quad \text{is the cross-bispectrum.}
\]

Control variates may be used to achieve further variance reduction. Suppose that we can find some simple function \( \hat{L}_0 \), say, of the past observations, which approximates reasonably closely the performance measure \( L_0 \) and whose expectation is analytically calculable and differentiable. For example, we might take \( \hat{L}_0 \) to be a linear combination of functions \( g_i(X_i) \) such that \( \frac{\partial}{\partial \nu} E_{\nu}(X_i) \)
can be calculated analytically; or simply
\[ \hat{L}_t = c(\nu) + \sum_{i=0}^{\infty} d_i(\nu) X_{t-i} \]  
(2.9)
for nonrandom regression coefficients \( c(\nu) \) and \( d_i(\nu) \); or in the case of a GI/G/1 queue with \( L_t \) the sojourn time of a customer, \( \hat{L}_t \) might be a weighted average of the difference between the service times and the interarrival times of a few of the preceding customers in the system. Then
\[
\frac{\partial}{\partial \nu} I(\nu) = \frac{\partial}{\partial \nu} \mathbb{E}_\nu[L_t-L_t'] + \frac{\partial}{\partial \nu} \mathbb{E}_\nu[\sum_{i=0}^{\infty} S_i] + \frac{\partial}{\partial \nu} \mathbb{E}_\nu L_t \]  
(2.10)
Thus we may estimate the cross spectral density function between \( L_t-L_t' \) and \( S_t \) rather than \( L_t \) and \( S_j \), the advantage being that judicious choice of \( \hat{L}_t \) may result in a cross spectral density function that is flatter near the origin (for example, a preliminary simulation may be used and \( L_t \) regressed on the preceding \( X_j \)). We may then use a spectral density estimator with wide window without substantially increasing the bias of the estimator while reducing its variance significantly.

For a regenerative process, there is some random time \( \tau \) such that \( L_t \) is independent of \( S_j ; j < t-\tau \). In this case \( t-\tau \) is a regeneration time of the process. Then since \( \mathbb{E}(L_tS_{t-k}) = \mathbb{E}(L_tS_{t-k} \mid k \leq \tau) F(k \leq \tau) \) we may estimate the cross-covariance using only terms \( L_t \) and \( S_j \) in the same regenerative cycle, further reducing the variance of the score function estimator to \( O(T^{-1}) \).

Further details will be given elsewhere.

ACKNOWLEDGMENTS. This research was supported by grants from the Natural Sciences and Engineering Research Council of Canada.
REFERENCES


Andrey Feuerverger  
Department of Statistics  
University of Toronto  
Toronto, Ontario  
Canada M5S 1A1

Don L. McLeish  
Department of Statistics  
University of Waterloo  
Waterloo, Ontario  
Canada N2L 3G1

Reuven Rubinstein  
Faculty of Industrial Engineering and Management  
The Technion  
Haifa 32000  
Israel

Received 6 August, 1986