Chapter 25

VARIATONAL STATIONARY FUNCTIONALS IN MONTE CARLO COMPUTATIONS

12/10/2013

INTRODUCTION

Use of stationary functionals from variational theory in Monte Carlo computations is discussed. Minimum variance secondary estimators based on the optimal linear combination of primary estimators for the terms constituting these functionals is suggested, and a two-stage procedure using these estimators is used to estimate inner products over discrete random walks. It is shown that the optimal combination yields econdary estimators with a variance at least equal to the variance of the used primary estimator with the least variance; so that gains from applying the method can be expected in conjunction with the use of stratified sampling, where different estimators in different regions of phase space have different variances [1].

Integral properties of physical systems can be mathematically formulated in terms of functionals depending upon functions satisfying the equations governing the system. If approximate solutions are readily available, and if a variational principle can be written for the property of interest, then this quantity can be evaluated with an error which is of second order with respect to the error in the approximate solution relative to the exact solution. This stationarity property suggested the exploration of these variational principles for variance reduction in Monte Carlo nuclear reactor computations [2]. Nakache, Kalos, Goldstein and Celnik [3, 4, 5], introduced a Monte Carlo method using these functionals, based on evaluating analytically or numerically some of the terms constituting these functionals.

In an attempt to generalize the Nakache et al. method for Monte Carlo applications by eliminating the need for analytical or numerical evaluation of some terms of the stationary functionals used, it was established that the statistical behavior of the estimation process must be taken into account. The condition of minimum variance in the Monte Carlo case replaces that of cancellation of first order errors in the deterministic case. Based on that consideration, Monte Carlo estimators are suggested in an approach using linear regression Monte carlo methods to optimally combine the terms of stationary functionals from variational theory. The usefulness of these estimators is apparent when combined with other standard variance reduction methods such as stratification of phase space. Particularly, a linear regression two-stage sequentialapproach yields a modification to the Nakache et al. method.

Regression methods applied to Monte Carlo were previously considered by several authors: Spanier and Gelbard et al. [6, 7, 8], Gelbard et al. [6, 7], MacMillan [9], Hammersley and Handscomb [10] and Halperin [20].

In this work, linear regression minimum variance Monte Carlo estimators related to stationary variational principles for source problems are suggested. They are combined with a two-stage sequential scheme. The relationship to previous work, particularly to the

Nakache et al. method is discussed, The application of these estimators to the evaluation of inner products over discrete random walks using the last event (last collision) estimator, and some numerical results are described.

Borgwaldt and Kalos [12] reported that it was found that the use of optimally combined estimators does not yield the theoretically expected gains in variance reduction. Our numerical results show that the linear regression Monte Carlo yield secondary estimators which have at least the variance of the primary estimator with the least variance. As such, gains in the application of the optimum combination of estimators can only be obtained in relation to stratified sampling, in case the primary estimators have unequal variances in the different stratified regions of phase space. Gains from a proposed Modified Nakache et al. method can be achieved when sufficient correlation exists between the primary estimators used.

THEORY

Problem Formulation

Consider the source problem for a physical system governed by the equation:

$$\Xi\phi = S \tag{1}$$

and its adjoint:

$$\Xi^{\dagger}\phi^{\dagger} = S^{\dagger} \tag{2}$$

where: Ξ is a linear operator,
ϕ is a probability distribution function (pdf),
S is an independent source.

In particular, $\phi(\underline{P})$ may satisfy the integral Boltzmann Transport Equation with $\underline{P}$ its phase space point $(\underline{r}, E, \underline{\Omega}, t)$ where $\underline{r}, E, \underline{\Omega}, t$ represent spatial position, energy, direction, and time, respectively.

Suppose it is required to estimate by Monte Carlo an integral property of the system such as the linear functional:

$$LF(\phi) = < S^{\dagger}, \phi > = \int_{\underline{P}} S^{\dagger}(\underline{P})\phi(\underline{P})d\underline{P} \tag{3}$$

where: $S^{\dagger}$ is a response function characteristic of the system,
(e. g. a reaction cross section),
$\underline{P}$ represents all independent variables.

The response function is equal to the functional derivative of LF, i. e.:

$$S^{\dagger}(\underline{P}) = LF'(\phi)\,. \tag{4}$$

We notice that the following relationships hold:

$$< S^{\dagger}, \phi >=< S, \phi^{\dagger} >=< \phi^{\dagger}, \Xi\phi >=< \Xi^{\dagger}\phi^{\dagger}, \phi > \tag{5}$$

Thus the linear functional 3 may be obtained by either the forward mode estimate $< S^{\dagger}, \phi >$ or the adjoint mode estimate $< S, \phi^{\dagger} >$, according to the problem at hand. This reciprocity property was introduced by Maynard [6, 14] to Monte Carlo computations.

Linear-Regression Variational Secondary Estimators

The Roussopoulos functional [15, 16], can be modified to the estimator:

$$E_{MR}(\phi^{\dagger}, \phi; V_0) = m_1 < S^{\dagger}, \phi > + m_2 < \phi^{\dagger}, S > + m_3 < \phi^{\dagger}, \Xi\phi > \tag{6}$$

where: m_1, m_2 and m_3 are the components of the row vector:

$$\underline{M} = \{m_1\ m_2\ m_3\} = [\underline{\mathrm{U}}^T\, V_0^{-1}\, \underline{\mathrm{U}}]^{-1}\, \underline{\mathrm{U}}^T\, V_0^{-1} \tag{7}$$

The second form corresponds to a minimum variance linear optimal combination of the inner product terms in the modified Roussopoulos functional E_{MR} as analyzed in relation to linear regression Monte Carlo by Hammersley and Handscomb [10]. Here we require that:

$$m_1 + m_2 + m_3 = 1 \tag{8}$$

To insure an unbiased estimate. The vector $\underline{\mathrm{U}}$ has unit components, $\underline{\mathrm{U}}^T$ is its transpose, V_0 is the sample estimate of the variance-covariance matrix V for the three inner products in the expression 6, V_0^{-1} is the inverse of V_0, implying the condition for the existence of a solution for $\underline{M}$:

$$\det V_0 = 0 \tag{9}$$

The proposed modified estimator has a nearly minimum variance. It is only nearly minimum since V_0 is used instead of the unknown exact V, but first order deviations in the neighborhood of a minimum of a function, only cause second order variations [9]. Further,

$$\mathrm{var}(\mathrm{E}_{MR}) = [\underline{\mathrm{U}}^T\, V_0^{-1}\, \underline{\mathrm{U}}]^{-1} \tag{10}$$

and it is an unbiased estimator of the required linear expectation [10], i. e.:

$$mean\{E_{MR}(\phi^{\dagger},\phi,V_0)\} = LF \ . \tag{11}$$

The different terms of expression 6 can all be simulated by one single basic estimator, e. g. the analog estimator, collision estimator, path length estimator, etc., or by a combination of different basic estimators for each term. Numerical experimentation shows that the optimal linear combination will give agreater weight to the primary estimator with the smallest variance in different regions of phase space, in case the estimates for the different terms are positively correlated, and this suggests the use of stratification to write the secondary estimator:

$$E_{MRS} = \sum_{j=1}^{k} \alpha_j(\underline{P}) E_{MR}[p_j(\{\rho\}_j)] \tag{12}$$

with:
$$\alpha_j(\underline{P}) = \frac{\mathrm{v}_j(\underline{P})}{\mathrm{v}(\underline{P})}, \sum_{j=1}^{k} \alpha_j = 1 \tag{13}$$

The $\alpha_j(\underline{P})$'s are the weights assigned to the different strata of volumes $\mathrm{v}_j(\underline{P})$ in the whole volume $\mathrm{v}(\underline{P})$ of phase space, p_j is a phase space point in stratum j, $\{\rho_j\}$ denotes a set of pseudo-random numbers that generate a point in stratum j. These random numbers are usually (but not necessarily) associated with an early part of the history, e. g. position, energy, direction for the source, or length of the first free flight.

If the primary estimators for the different terms of E_{MR} possess different variances in different strata, the optimum combination will give a larger weight to terms with the least variance in each stratum. Consequently, E_{MRS} will yield an estimate with a smaller variance over the whole phase space. In case a negative correlation exists between the different terms in the expression for E_{MR}; a case closely related to antithetic variates would arise with an associated variance reduction, and stratification may not be needed.

An example of methods for creating a negative correlation is exposed by Spanier and Gelbard [6] in relation to the calculation of resonance integrals. There, one uses a pseudo-random number r to determine a neutron starting energy E, and the number (1-r) to determine a second starting energy E'. When r is close to 0 or 1, E and E' will have a tendency to be at opposite ends of the energy spectrum of the source, with subsequent negatively correlated contributions to the resonance integral for the paired particles.

It is interesting to notice that the linearly optimal combination of the natural logarithms of the inner products in expression 6, i. e.:

$$\ln[\mathrm{LF}(\phi^{\dagger},\phi)] = m_1' \ln < S^{\dagger},\phi > + m_2' \ln < \phi^{\dagger}, S > + m_3' \ln < \phi^{\dagger}, \Xi\phi > \tag{14}$$

yields a modified minimum variance optimal Swinger (fractional) estimator:

$$E_{MS}[\phi^{\dagger},\phi;V_0)] = \frac{<S^{\dagger},\phi>^{m_1'} . <\phi^{\dagger},S>^{m_2'}}{<\phi^{\dagger},\Xi\phi>^{m_3'}} \tag{15}$$

where: m_1, m_2 and m_3 are the components of the row vector:

$$\underline{M}' = \{m_1' m_2' m_3'\} = [\underline{\mathrm{U}}^T V_0^{-1} \underline{\mathrm{U}}]^{-1} \underline{\mathrm{U}}^T V_0^{-1}$$

has the same form as in Eqn. 7.

These estimators E_{MR} and E_{MS}, for the special choices:

$$m_1 = m_1' = m_2 = m_2' = 1,\ m_3 = m_3' = -1,$$

Reduces to the Roussopoulos and Swinger variational principles corresponding to the deterministic non-statistical case.

It is important to notice that the terminology used here is adopted from previous work [15, 16] and the E_{MR} and E_{MS} estimators do not imply second order errors, and first order error cancellation as in the deterministic (non-statistical) case. In fact, minimum variance in the statistical case replaces the corresponding first order error cancellation in the deterministic case. The statistical nature of the Monte Carlo process and the existence of correlations preclude the possibility of first order error cancellation. Zero variance could only be achieved with a delta function as the frequency function.

Clearly one needs to obtain sampling estimates of $\underline{M}$ and $\underline{M}'$, in order to use these estimators, and a two-stage process is to be used as described below.

Various generalizations of the quantity of interest to be estimated can be made in analogy to other variational functionals. These include the cases where one seeks the ratio of linear functionals, bilinear functionals, etc. [15,16]. The statistical analogues of these are easily stated.

MODIFIED NAKACHE-KALOS-GOLDSTEIN-CELNIK METHOD

Following the same approach of linear regression adopted in the previous section, let us consider the estimator:

$$E_{MN}[\phi^{\dagger},\phi;a,b] = t_1 + bt_2 + a \tag{16}$$

where t_1 and t_2 may be any two different functionals from those in Eqn. 5. For E_{MN} to be an unbiased estimator of the linear expectation μ, we require that:

$$mean[t_1 + bt_2 + a] = \mu_1 + b\mu_2 + a = \mu \tag{17}$$

Further, the expectations μ_1 and μ_2 of t_1 and t_2 must also equal μ, implying that:

$$\mu + b\mu + a = \mu, \Rightarrow a = -b\mu$$

Thus, Eqn. 16 can be rewritten as:

$$E_{MN}[\phi^{\dagger}, \phi; b] = t_1 + b(t_2 - \mu) \tag{18}$$

The variance of E_{MN} can be written as:

$$\begin{aligned} \mathrm{var}[E_{MN}] &= mean[\{t_1 + b(t_2 - \mu)\} - \mu]^2 \\ &= \mathrm{v}_{11} + b^2 \mathrm{v}_{22} + 2br_{12}\mathrm{v}_{11}^{1/2}\mathrm{v}_{22}^{1/2} \end{aligned} \tag{19}$$

where r_{12} is the linear correlation coefficient between t_1 and t_2.

For a minimum of the variance relative to b:

$$\frac{\partial\, \mathrm{var}[E_{MN}]}{\partial b} = 2b\mathrm{v}_{22} + 2r_{12}\mathrm{v}_{11}^{1/2}\mathrm{v}_{22}^{1/2} = 0$$

must be equal to zero, which implies:

$$b = -r_{12}\left(\frac{\mathrm{v}_{11}}{\mathrm{v}_{22}}\right)^{1/2} \tag{20}$$

Substituting from Eqn, 20 into Eqn. 19, we get the expression for the minimum variance:

$$\mathrm{var}[E_{MN}]_{\min} = \mathrm{v}_{11}(1 - r_{12}^2) \tag{21}$$

In analogy to Eqn, 15, we may also write:

$$E'_{MN}[\phi^{\dagger}, \phi; c] = \frac{t_1.\mu}{t_2^{-c}} \tag{22}$$

where "c" has the same form as "b" in Eqn. 20.

Clearly, better variance reduction will be obtained for higher positive or negative linear correlation between the chosen t_1 and t_2. Again a two-stage process is a necessity to evaluate sample estimates of μ and b (or c) by Monte Carlo. Ideally, for complete positive correlation ($r_{12} = +1$), Eqn. 21 predicts an ideal zero-variance sampling scheme.

ESTIMATION OF INNER PRODUCTS OVER TERMINATING DISCRETE RANDOM WALKS

Problem Formulation

The Monte Carlo treatment of Transport Theory often employs a discrete treatment of the energy variable coupled to a continuous treatment of the spatrial variables. This justifies the use of discrete random walks and linear systems to deduce estimators and numerically check the theoretical work.

We consider Eqn. 1 with:

$$\Xi = I - H,$$

in the form:

$$\underline{\phi} = H\underline{\phi} + \underline{S} \tag{23}$$

which describes the N-group particle transport process in an infinite homogeneous, non-multiplying medium, with a constant isotropic source $\underline{S}$ [6].

Here: H is the transition matrix, whose entries h_{ij} may represent the transition probabilities from energy group j to energy group i ,

$\underline{S}$ is the density of births, or density of first collisions, assumed normalized,

$\underline{\phi}$ is the collision density whose j -th component, ϕ_j represents the expected number of collisions in energy group , per source particle.

For convergence, the spectral radius of H should satisfy the condition:

$$\rho(H) < 1,$$

or equivalently:

$$\max_r \left|\lambda_r(H)\right| < 1$$

where $\lambda_i(H)$ are the eigen-values of H .

On physical grounds:

$$\phi_i \geq 0, S_i \geq 0, \forall i,$$

and:

$$h_{ij} \geq 0, \forall i, \mathrm{j}.$$

Deduction of Primary Estimators

A discrete random walk process α is defined by a set of first collision probabilities in state i : p_i^1, a set of transmission probabilities p_{ij} from state j to state i, and a set of death or absorption probabilities in state i: p_i, such that [6]:

$$\begin{aligned} w_{ij} &= \frac{h_{ij}}{p_{ij}}, if\ p_{ij} \neq 0 \\ &= 0 \quad , if\ p_{ij} = 0 \end{aligned} \tag{24}$$

are statistical weight factors which can be used to simulate the linear problem.

Let:

$$\alpha = (i_1, i_2, \ldots, \mathrm{i}_k)$$

be an arbitrary random walk starting in state i_1 and terminating at state i_k. We are interested in deducing primary estimators for the terms in expression 5 over a terminating discrete random walk model.

THEOREM I

The "last collision" or "von Neumann-Ulam-Forsyth-Leibler [17]" unbiased estimator of the inner product $< S^\dagger, \phi >$ over the discrete random walk α is:

$$\xi_a(\alpha) = \frac{S_{i_1}}{p_{i_1}^1} \left(\prod_{l=1}^{k-1} w_{i_{l+1} i_l} \right) \frac{S_{i_k}^\dagger}{p_{i_k}} \tag{25}$$

A proof is carried out by inserting the Neumann series for ϕ in $< S^\dagger, \phi >$ and substituting for the h_{ij}'s by $w_{ij} p_{ij}$ from Eqn. 24. The result is:

$$\sum_{k=1}^{\infty} \sum_{i_k} \sum_{i_1} \frac{S_{i_1}}{p_{i_1}^1} \left(\prod_{l=1}^{k-1} w_{i_{l+1} i_l} \right) p_{i_1}^1 \left(\prod_{l=1}^{k-1} p_{i_{l+1} i_l} \right) \frac{S_{i_k}^\dagger}{p_{i_k}^1}$$

Which is recognized as the sum over $\xi_a(\alpha) P(\alpha)$,

$$< S^\dagger, \phi > = \sum_{\alpha} \xi_a(\alpha) P(\alpha)$$

where $P(\alpha)$ is the chain probability:

$$P(\alpha) = p_{i_1}^1 p_{i_2 i_1} \ \ldots \ p_{i_k i_{k-1}} p_{i_k}$$

THEOREM II

The "last collision" estimator of the inner product $< S, \phi^{\dagger} >$ over the random walk α is:

$$\xi_b(\alpha) = \frac{S^{\dagger}_{i_1}}{p^1_{i_1}} \left(\prod_{l=1}^{k-1} w_{i_{l+1} i_l} \cdot \frac{h_{i_l i_{l+1}}}{h_{i_{l+1} i_l}} \right) \frac{S_{i_k}}{p_{i_k}} \tag{26}$$

The proof proceeds by replacing $\phi^{\dagger}$ by its Neumann series in $< S, \phi^{\dagger} >$. The term for k collisions is then multiplied and divided by the factor:

$$p^1_{i_1} \left(\prod_{l=1}^{k-1} p_{i_{l+1} i_l} \right) p_{i_k}$$

defining the probability of a track α of k collisions.

Replacing the p_{ij}'s by $\frac{h_{ij}}{w_{ij}}$, $h^{\dagger}_{ij}$ by h_{ji}, one recognizes the result:

$$< S, \phi^{\dagger} > = \sum_{\alpha} \xi_b(\alpha) P(\alpha)$$

Here $\xi_b(\alpha)$ is written to show how it is related to $\xi_a(\alpha)$ by corrective factors, so that both ξ_a and ξ_b can be estimated over the same track α. Notice that our summation of the Neumann series is over all track lengths from 1 to infinity.

REFERENCES

1. Magdi M. H. Ragheb, "Investigation of New Estimation Approaches for Nuclear Reactor Computations by Monte Carlo," Ph. D. Thesis, University of Wisconsin, Madison, Wisconsin, USA, 1978.
2. M. H. Kalos, "Adjoint Monte Carlo – What is it? Who needs it?," in: "A Review of the Monte Carlo Method for Radiation Transport," B. T. Maskewitz and V. Z. Jacobs, Eds., ORNL-RSIC-29, 1971.
3 .F. R. Nakache, "Variational Monte Carlo Calculations of Thermal Utilization Factors and Neutron Depression in Foils, Thesis Columbia University, 1962.
4. F. R. Nakache, M. H. Kalos and H. Goldstein, "Variational Monte Carlo Calculations of Flux Depression in Foils," Trans. Am. Nucl. Soc., 5, 1, p. 32, 1962.
5. M. H. Kalos, F. R. Nakache and J. Celnik, "Monte Carlo Methods in Reactor Computations," in: "Computing Methods in Reactor Physics," H. Gfreenspan, C. N. Kelber and D. Okrent, Eds., Gordon Breach Science Publishers Inc. New York, 1968.
6. J. Spanier and E, M, Gelbard, "Monte Carlo principles and Neutron Transport Problems," Addison-Wesley Publishing Company, Reading, Massachusetts, 1969.

7. E. M. Gelbard, L. A. Ondis II and J. Spanier, " A New Class of Monte Carlo Estimators," J. SIAM Appl. Math., V. 14, p. 697, 1966.
8. J. Spanier, "Two pairs of Families of Estimators for Transport Problems," J. SIAM Appl. Math., Vol. 14, No.4, p.710, 1966.
9. D. B. MacMillan, "Comparison of Statistical Estimators for Neutron Monte carlo Calculations," Nucl. Sci. Eng., 10, pp 97-101, 1961.
10. J. M. Hammersley and D. C. Handscomb, "Monte Carlo Methods," Methuen and Co. Ltd, London, 1967.
11. A. W. Marshall, "The Use of Multistage Sampling Schemes in Monte Carlo Methods," in: "University of Florida Symposium on Monte Carlo Methods," H. A. Meyer, Ed., New York, Wiley, pp. 37-46, 1956.
12. Panel Discussions, "Proceedings of the NEACRP Meeting of a Monte Carlo Study Group," ANL-75-2, NEA-CRP-L-118, p. 238, 1974.
13. L. L. Carter and E. D. Cashwell, "Particle Transport Simulation with the Monte Carlo Method," ERDA Critical Review Series, TID-26607, 1975.
14. C. W. Maynard, " An Application of the Reciprocity Theorem to the Acceleration of Monte Carlo Computaions, Nucl. Sci. Eng., 10, pp. 97-101, 1961.
15. W. M. Stacey, Jr., "Variational Methods in Nuclear Reactor Physics," Academic Press, p. 6, 1974.
16. R. Conn and W. M. Stacey, Jr., "Variational Methods for Controlled Thermonuclear Reactor Blanket Studies," Nuclear Fusion, 13, p. 185, 1973.
17. G. E. Forsyth and R. A. Leibler, "Matrix Inversion by a Monte Carlo Method," MTAC, V. 4, pp.127-129, 1950.
18. W. R. Wasow, "A Note on the Inversion of Matrices by Random Walks," MTAC, V. 6, pp.78-81, 1952.
19. M. J. D. Powell and J. Swann, "Weighted Uniform Sampling – A Monte Carlo Technique for Reducing Variance," J. Inst. Maths. Applics 2, pp. 228-236. 1966.
20. M. Halperin, "Almost Linearly Optimum Combination of Unibiased Estimates," J. Amer. Statist. Assoc., 56, p. 36, 1961.