Extension of the hybrid Monte Carlo method for boundary-value problems

by

Ludwig Fahrmeir
Institut für Angewandte Mathematik
Technische Universitat München
D-8000 München 2, Arcisstraße, Germany

BIOGRAPHY

After a diploma in mathematics and physics from the Technical University of Munich in 1970, LUDWIG FAHRMEIR accepted an offer to work at the Institut für Angewandte Mathematik as a member of the hybrid computing systems research group, whose two principal aims were to develop a hybrid computing system interfacing a Beckman Ease analog and a CAE 90-10 digital computer, and to develop hybrid computing methods to solve problems in applied mathematics. In the course of this work, Dr. Fahrmeir investigated hybrid methods for solving boundary-value problems, coming up with the extension of the hybrid Monte Carlo method described in this paper. In December 1972 he completed his dissertation and is now working as an assistant at the Institut. His research interests embrace the theory and application of stochastic processes, including further development of Monte Carlo methods. He lives with his wife in Germering, near Munich, where he enjoys skiing, sailing, and guitar playing.

ABSTRACT

Hybrid Monte Carlo techniques for the solution of linear boundary-value problems have previously been developed. This paper is not primarily concerned with implementing these techniques, but extends the class of problems that can be solved by them and improves the method first described by Little, which in its original form is shown to be valid only in a special case and not generally applicable to linear boundary-value problems.

INTRODUCTION

A hybrid Monte Carlo method for solving certain classes of linear boundary-value problems was first developed by Chuang, Kazda, and Windeknecht[2]. Little[10] described a generalization of this technique to yield a method for solving the Dirichlet problem for elliptical partial differential equations with "slowly varying" coefficients and without mixed second-space derivatives. Handler[6] extended it to other boundary conditions, employing the Astrac II ultrahigh-speed computer at the University of Arizona.[9]

The hybrid Monte Carlo technique is also described by Bekey and Karplus[1] and Korn.[9] I investigated the hybrid Monte Carlo method by a different mathematical approach[5] using the relation between Markov processes and partial differential equations,[4] extending this method to general linear elliptic equations, and demonstrating that Little's method is only valid in the special case of constant coefficients of the second derivatives. Generally the two methods differ in the simulation of the random walks in certain stochastic differential equations. In the special case the two methods coincide. The other aspects of computer implementation[3,5,7,8,9,10]--such as noise generation, detection of boundary crossings, error analysis, and computation time--are not affected by my extension.

BOUNDARY-VALUE PROBLEMS AND SOLUTION BY THE MONTE CARLO METHOD

I shall describe the method briefly, omitting its complete mathematical derivation, which is in my dissertation. However, the approach to the mathematical derivation is suggested in the remark at the end of this section.

Consider the following boundary-value problem (the Dirichlet problem) for the real function $u(x) = u(x_1,\ldots,x_n)$ on a domain G of n-dimensional Euclidean space with boundary G' $u(x)$ is required to satisfy the elliptic differential equation:

$$\sum_{i=1}^{n} a_i(x)\,\frac{\partial u(x)}{\partial x_i} + \frac{1}{2}\sum_{i,j=1}^{n} b_{ij}(x)\,\frac{\partial^2 u}{\partial x_i \partial x_j} - f(x)u(x) = g(x) \qquad (1)$$

where x is in G and where the boundary condition

$$\lim_{x \to r} u(x) = \psi(r) \qquad (2)$$

where $r = (r_1,\ldots,r_n)$, lies on G' and $\psi(r)$ is a continuous function on G'.

If the domain G and its boundary G' satisfy certain regularity conditions and

If the coefficients b_{ij} possess continuous first derivatives, and (3)

If these derivatives and the coefficients a_i, f, g satisfy on $G \cup G'$ a Hölder condition with exponent λ (e.g., $|f(x) - f(y)| \le L||x - y||^{\lambda}$ $||\ ||$ being the distance between x and y, with L a positive constant), and (4)

If for any x in $G \cup G'$ and any collection of nonzero real numbers $l_1,\ldots,l_n$ (5)

$$\sum_{i,j=1}^{n} b_{ij}(x)\, l_i\, l_j \ge c \sum_{i=1}^{n} l_i^2, \quad c > 0$$

If $f(x) \ge 0$ in G (6)

then[4,11] there exists a unique solution $u(x)$ of the Dirichlet problem (1), (2).

Let $B(x)$ be the matrix with elements $b_{ij}(x)$, assuming that $B(x)$ is symmetric. $\overline{C}(x)$ denotes the symmetric matrix satisfying

$$\overline{C}^2(x) = B(x) \qquad (7)$$

Condition (5) guarantees the existence of $\bar{C}(x)$, but it may be difficult to obtain it if $n \geq 3$ and $B(x)$ has many off-diagonal elements. It is easy if (as assumed by Little *et al.*) $B(x)$ is a diagonal matrix. Then

$$\bar{c}_{ii}(x) = \sqrt{b_{ii}(x)} \text{ and } \bar{c}_{ij}(x) = 0$$

if $i \neq j$ ($\bar{c}_{ij}$ elements of $\bar{C}$).

Now let $N_i(t)$ be n independent noise sources approximating Gaussian white noise with power spectral density σ_i^2, suitable for application to high-speed Monte Carlo methods.[3,6,7,8,9] S denotes the diagonal matrix with elements σ_i, and $C(x)$ denotes the matrix defined by

$$C(x) = \bar{C}(x) \cdot S^{-1} \tag{8}$$

where S^{-1} is the inverse of S.

If $B(x)$ is diagonal, (7) and (8) result in

$$c_{ii}^2(x) = b_{ii}(x)/\sigma_i^2$$

This is the formula of Little *et al.*

Suppose now that we desire the solution $u(x)$ of the boundary-value problem at $\bar{x} = (\bar{x}_1,\ldots,\bar{x}_n)$ somewhere in G. Then a n-dimensional random walk $X(t) = (X_1(t),\ldots,X_n(t))$ can be initiated from point $\bar{x}$ by integrating the following system of stochastic differential equations

$$\frac{dX_i(t)}{dt} = a_i(X(t)) - \frac{1}{2}\sum_{j,k=1}^{n} \bar{c}_{kj}(X(t)) \frac{\partial \bar{c}_{ij}(X(t))}{\partial x_k} + \sum_{j=1}^{n} c_{ij}(X(t))N_i(t); \quad i = 1,\ldots,n \tag{9}$$

with initial condition

$$X_i(0) = \bar{x}_i$$

If τ is the first exit time from G after $t = 0$, then the value of the boundary condition at the point of crossing is denoted by $\psi(X(\tau))$. With τ and $\psi(X(\tau))$ we define

$$\Phi(X(\tau)) = \exp\left(-\int_0^\tau f(X(t))dt\right)\psi(X(\tau)) - \int_0^\tau \exp\left(-\int_0^\tau f(X(s))ds\right) g(X(t))dt \tag{10}$$

If $f(x) \equiv g(x) \equiv 0$ in G, then Φ reduces to ψ.

We start now a number N of random walks and denote by Φ_i the value of $\Phi(X(\tau))$ realized by the ith random walk. The average

$$\bar{\Phi} = \frac{1}{N}\sum_{i=1}^{N} \Phi_i \tag{11}$$

gives an approximation of $u(\bar{x})$:

$$u(\bar{x}) \approx \bar{\Phi}$$

COMPARISON WITH LITTLE'S METHOD

Compare now this Monte Carlo method (MC1) with the method first described by Little (MC2):

In MC2 the coefficients $b_{ij}(x)$ are restricted to $b_{ij}(x) \equiv 0$ if $i \neq j$. The random walk $\overset{*}{X}(t) = (\overset{*}{X}_1(t),\ldots,\overset{*}{X}_n(t))$ is governed by

$$\frac{d\overset{*}{X}_i(t)}{dt} = a_i(\overset{*}{X}(t)) + c_{ii}(\overset{*}{X}(t))N_i(t) \tag{12}$$

$$\overset{*}{X}_i(0) = \bar{x}_i; \quad i = 1,\ldots,n$$

Then MC2 continues as MC1 with $\overset{*}{X}(t)$ instead of $X(t)$.

Comparing (9) and (11), we see that

1) MC1 is applicable to a wider class of boundary-value problems, since the coefficients $b_{ij}(x)$, $i \neq j$, are not restricted to $b_{ij}(x) \equiv 0$

2) even when $b_{ij}(x) \equiv 0$, $i \neq j$, MC1 differs from MC2 by the additional terms $-\frac{1}{2}\sum_{j,k=1}^{n} \bar{c}_{kj}\frac{\partial \bar{c}_{ij}}{\partial x_k}$ in (9).

Only for the case of a constant diagonal matrix $B(x) = B$ does MC1 reduce to MC2.

Little assumed that the coefficients $b_{ii}(x)$ were "slowly varying." The additional terms then may be small if the derivatives $\partial \bar{c}_{ij}/\partial x_k$ are small enough. Generally, however, these additional terms are not negligible.

Remark: The absence of the additional terms in MC2 is a consequence of Little's derivation of the method. He writes (12) assuming $N_i(t)$ to be ideally Gaussian white noise. But white noise is not integrable (with probability 1) and therefore (12) has no mathematical meaning. By integrating (12), Little gets the (mathematically undefined) integral equation

$$\overset{*}{X}_i(t) = \bar{x}_i + \int_0^t a_i(\overset{*}{X}(s))ds + \int_0^t c_{ii}(\overset{*}{X}(s))N_i(s)ds \tag{13}$$

Then he--intuitively--applies the rules of the Ito-calculus.[4] If we write (13) in the form of Ito's stochastic integrals

$$\overset{*}{X}_i(t) = \bar{x}_i + \int_0^t a_i(\overset{*}{X}(s))ds + \int_0^t c_{ii}(\overset{*}{X}(s))dw_i(s) \tag{14}$$

the rest of Little's proof, in the main, remains unchanged. A more general result, where the $c_{ij}(x)$ are not restricted to be zero, may be found in Dynkin's book.[4] But the realizations of $\overset{*}{X}_i(t)$ cannot be obtained directly by (14) because Ito's stochastic integrals are defined on the quadratic mean.[5] If the physically realisable noise $N_i(t)$ in (9) converges to white noise, the process $X_i(t)$ defined by (9) converges to $\overset{*}{X}_i(t)$ as defined by (14). If additionally $N \to \infty$, the average $\bar{\Phi}$ converges to $u(\bar{x})$. The additional terms in (9) arise from the unsymmetric definition of the second integral in (14). The proofs of these statements are given in my dissertation.[5]

EXAMPLE

The example given in this paper is selected from a number of problems[5] which were solved using the hybrid computing system (Beckman Ease 2133/CAE-90-10) of the Institut für Angewandte Mathematik at the Technische Universität München, which is directed by Professor J. Heinhold. The computer implementation differs from Little's only by the simulation of the stochastic differential equations. Noise generators with $\sigma = \sigma_1 = \sigma_2 = 0.096$ were used.

Boundary value problem

$$\frac{\sigma^2}{2}\{-4x\frac{\partial u}{\partial x} + (x^2+y^2)\frac{\partial^2 u}{\partial y^2} + 4yx\frac{\partial^2 u}{\partial x \partial y} + (x^2+y^2)\frac{\partial^2 u}{\partial^2 y}\} = 0$$

with boundary condition (see Figure 1):

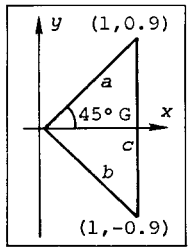

Figure 1.

$$\psi(x,y) = \begin{cases} x^3 - \frac{1}{3}(x-0.1)^3 - 0.1x^2 + 1, & (x,y)\varepsilon a \\ \frac{1}{3}(x-0.1)^3 - x^3 + 0.1x^2 + 1, & (x,y)\varepsilon b \\ y - \frac{1}{3}y^3 + 1, & (x,y)\varepsilon c \end{cases}$$

Exact solution

$$u(x,y) = x^2 y - \frac{1}{3}y^3 + 1$$

With $B = \sigma^2 \begin{pmatrix} x^2+y^2 & 2yx \\ 2yx & x^2+y^2 \end{pmatrix}$ and $S = \begin{pmatrix} \sigma & 0 \\ 0 & \sigma \end{pmatrix}$,

we get $\overline{C} = \sigma \begin{pmatrix} y & x \\ x & y \end{pmatrix}$ and $C = \begin{pmatrix} y & x \\ x & y \end{pmatrix}$

By (9) the random walks (N = 1000) are governed by the stochastic differential equations

$$\frac{dX(t)}{dt} = -3\sigma^2 X(t) + Y(t)N_1(t) + X(t)N_2(t)$$

$$\frac{dY(t)}{dt} = -\sigma^2 Y(t) + X(t)N_1(t) + Y(t)N_2(t)$$

Results

For each (x,y), the value in the first line is computed by the Monte Carlo method, the second value is exact):

y \ x	.4	.5	.6	.7	
.4			1.128	1.149	MC
			1.123	1.175	exact
.3		1.085	1.091	1.118	MC
		1.066	1.100	1.144	exact
.2	1.045	1.049	1.065	1.070	MC
	1.031	1.047	1.069	1.096	exact
.1	1.018	1.032	1.044	1.044	MC
	1.016	1.025	1.036	1.049	exact
.0	1.009	.999	1.011	1.005	MC
	1.000	1.000	1.000	1.000	exact
-.2		.957	.952	.931	MC
		.953	.930	.905	exact
-.4			.887	.857	MC
			.877	.804	exact

CONCLUSION

I have described a Monte Carlo method for obtaining approximate solutions of the first boundary-value problem for linear elliptical differential equations which corrects and extends the method of Little.

He applied his method also to parabolic differential equations with known boundary conditions and initial conditions, and Handler extended it to other boundary conditions (e.g., the Neumann problem). Although I have yet to test it numerically, I am quite sure that the correction and extension given in my method also

applies to these problems, since the stochastic differential equations which govern the random walks depend only on the coefficients of the first and second derivatives of the partial differential equation and not on the boundary conditions.

REFERENCES

1 BEKEY G KARPLUS W
Hybrid Computation
Wiley New York 1968 pp 240-243

2 CHUANG K KAZDA L WINDEKNECHT T
A Stochastic Method of Solving Partial Differential Equations Using an Electronic Analog Computer
Project Michigan report 2900-91-T Willow Run Laboratories University of Michigan June 1960

3 CIMORELLI L CAPARASO R FEDERICO A
Hybrid Stochastic Monte Carlo Methods for Solving Heat Transfer Problems in Nuclear Reactors
AICA Annals 1968 pp 195-205

4 DYNKIN E B
Markov Processes vol 2
Springer Verlag Berlin 1965 pp 46-51

5 FAHRMEIR L
Schwache Konvergenz von Wahrscheinlichkeitsmaßen und ein Hybrides Monte-Carlo-Verfahren für Lineare Randwertaufgaben
Dissertation Institut für Angewandte Mathematik Technische Universität München, Germany 1972

6 HANDLER H
High Speed Monte Carlo Technique for Hybrid Computer Solution of Partial Differential Equations
PhD dissertation Department of Electrical Engineering University of Arizona 1967

7 JERMANN W CALHOUN M THOMAS R
Hybrid Monte Carlo Techniques with a Minimal Interface
SIMULATION vol 17 no 6 1971 pp 225-233

8 JOHNSON E
A Variance Reduction Technique for Hybrid Computer Generated Random Walk Solution of Partial Differential Equations
Proceedings Spring Joint Computer Conference
1970 pp 19-29

9 KORN G
Random-Process Simulation and Measurements
McGraw-Hill New York 1966

10 LITTLE W D
Hybrid Computer Solutions of Partial Differential Equations by Monte Carlo Methods
PhD thesis University of British Columbia 1965

11 MIRANDA C
Partial Differential Equations of Elliptic Type
Springer-Verlag Berlin 1970 pp 67-75