Numerical solution of boundary value problems for stochastic differential equations on the basis of the Gibbs sampler [1]

S.M. Prigarin [2] and G. Winkler [3]

To solve boundary value problems for linear systems of stochastic differential equations we propose and justify a numerical method based on the Gibbs sampler. In contrast to the technique which yields for linear systems an "exact" numerical solution, the proposed method is simpler to generalize for stochastic partial differential equations and nonlinear systems. Such generalizations are discussed as well.

1 Statement of the problem for linear stochastic differential equations

Let us consider a boundary value problem for a linear vector-valued stochastic differential equation (SDE) of the Ito type

$$du(t) = A(t)u(t)dt + \Sigma(t)dw(t), \quad t \in (t_1, t_N), \tag{1}$$

$$u(t_1) = u_*, \quad u(t_N) = u^*, \tag{2}$$

where $u(t) = (u_1(t), \ldots, u_k(t))^T$ is a vector-valued random process, $w(t) = (w_1(t), \ldots, w_k(t))^T$ is a vector of independent standard Wiener processes, $A(t)$ and $\Sigma(t)$ are $k \times k$ matrices whose elements are piecewise continuous functions, u_* and u^* are non-random vectors.

The problem is to construct realizations of the process $u(t)$ satisfying SDE (1) and boundary conditions (2). In other words, it is necessary to simulate a set $U = \{u(t_1), u(t_2), \ldots, u(t_N)\}$ of dependent random vectors $u(t_i)$ (for a fixed grid $T = \{t_1, t_2, \ldots, t_N\}$) with a joint distribution $P(U)$ generated by the SDE and the boundary conditions.

---

[1] The work partially was supported by RFBR grant N 00-15-96173

[2] Institute of Computational Mathematics and Mathematical Geophysics, pr. Lavrentieva 6, Novosibirsk, 630090, Russia; e-mail: smp@osmf.sscc.ru

[3] Institute of Biomathematics and Biometry, GSF-National Research Center for Environment and Health, 85764 Neuherberg München, Germany; e-mail: gwinkler@gsf.de

2 "Exact" numerical solution of boundary value problems for systems of linear stochastic differential equations

For a solution $u(t)$ of the linear SDE (1) the finite-dimensional distributions can be described in a comparatively simple way: if the values $u(\alpha)$ and $u(\beta)$ are fixed, then the distribution of $u(\gamma)$, $\gamma \in (\alpha, \beta)$, is Gaussian and does not depend on the values of $u(s)$ for $s < \alpha$ and $s > \beta$. The moments of the Gaussian distribution are written down in the following two statements (for details and proofs see [1,2]).

Here by $\Phi(t,s)$ we denote the principal matrix associated with $A(t)$ and by R_s^t we denote Gram's matrix

$$R_s^t = \int_s^t \Phi(t,\tau)\Sigma(\tau)\Sigma^T(\tau)\Phi^T(t,\tau)d\tau.$$

Lemma 2.1 (conditional distributions for linear SDE) *Assume that a vector-valued random process $u(t)$ satisfies SDE (1) on the segment $[\alpha, \beta]$, and the matrix R_α^β is nonsingular. Then the distribution of the vector $u(\gamma)$, $\gamma \in (\alpha, \beta)$, provided that $u(\alpha) = u_\alpha$, $u(\beta) = u_\beta$ (the vectors u_α, u_β are non-random), is Gaussian with expectation*

$$\begin{aligned} m &= \left[\Phi(\gamma,\alpha) - R_\alpha^\gamma \Phi^T(\beta,\gamma)(R_\alpha^\beta)^{-1}\Phi(\beta,\alpha)\right] u_\alpha + \\ &\quad + R_\alpha^\gamma \Phi^T(\beta,\gamma)(R_\alpha^\beta)^{-1} u_\beta \end{aligned} \tag{3}$$

and correlation matrix

$$\begin{aligned} &\mathbf{E}\,(u(\gamma) - m)\left(u^T(\gamma) - m^T\right) = \\ &= R_\alpha^\gamma - R_\alpha^\gamma \Phi^T(\beta,\gamma)(R_\alpha^\beta)^{-1}\Phi(\beta,\gamma)R_\alpha^\gamma. \end{aligned} \tag{4}$$

If the matrix R_α^β is singular and $u_\beta \in \Phi(\beta,\alpha)u(\alpha) + R_\alpha^\beta(\mathbf{R}^k)$, then formulas (3), (4) remain valid with pseudo-inverse matrix $(R_\alpha^\beta)^+$ instead of $(R_\alpha^\beta)^{-1}$.

Lemma 2.2 (conditional distributions for time-invariant linear SDE) *Assume that the vector-valued random process $u(t)$ is a solution to SDE (1) with constant matrices, and the matrix A is stable (i.e., $A(t) = A$ and $\Sigma(t) = \Sigma$ do not depend on t and the real parts of the roots λ of the characteristic equation $det(A - \lambda I) = 0$ are strictly negative). Then the distribution of the vector $u(\gamma)$, $\gamma \in (\alpha, \beta)$, provided that $u(\alpha) = u_\alpha$, $u(\beta) = u_\beta$, is Gaussian with expectation*

$$m^* = (B_1 - K_1 B_2^T K^{-1} B)u_\alpha + K_1 B_2^T K^{-1} u_\beta$$

and correlation matrix

$$R^* = \mathbf{E}\left(u(\gamma) - m^*\right)\left(u(\gamma) - m^*\right)^T = K_1 - K_1 B_2^T K^{-1} B_2 K_1,$$

where $B_1 = \exp((\gamma - \alpha)A)$, $B_2 = \exp((\beta - \gamma)A)$, $B = B_1 B_2 = B_2 B_1$,

$$\begin{aligned} &K_1 A^T + A K_1 = B_1 \Sigma \Sigma^T B_1^T - \Sigma \Sigma^T, \\ &K_2 A^T + A K_2 = B_2 \Sigma \Sigma^T B_2^T - \Sigma \Sigma^T, \\ &K A^T + A K = B \Sigma \Sigma^T B^T - \Sigma \Sigma^T, \\ &K = B_2 K_1 B_2^T + K_2. \end{aligned}$$

(Here we assume that the matrix K is nonsingular. In particular, K is nonsingular if the matrix Σ is nonsingular.)

According to the "exact" algorithm, the grid T is sequentially scanned in a certain order. For the current grid point $\gamma \in T$, the closest points $\alpha < \gamma$ and $\beta > \gamma$ must be found, where the values of u are already known, (note, that at the very beginning only the values $u(t_1)$, $u(t_N)$ are known). Then the value $u(\gamma)$ is simulated according to the conditional distribution described in Lemma 1. (Calculation of averages and correlation matrices for the conditional distributions in Lemma 1 is one of the most labor-consuming steps of the algorithm.) Finite-dimensional distributions of the process $u(t)$ are reproduced exactly by this algorithm.

There are no restrictions for the order of scanning, but computational costs can appreciably depend on it. From this point of view the bisection scheme seems to be efficient for time-invariant equations.

Further results and examples concerning the "exact" algorithm, including the study of existence and uniqueness of the solution for more general boundary conditions, can be found in [2,3].

3 The Gibbs sampler to solve boundary value problems for linear systems of SDE

For the method on the basis of the Gibbs sampler, we propose below, it is necessary to find conditional distributions of a special kind. For all inner grid points $t_i \in \{t_2, \ldots, t_{N-1}\} = T_0$ it is necessary to know the distributions of $u(t_i)$ provided that the values $u(t_{i-1})$, $u(t_{i+1})$ are fixed. These distributions will be denoted by $P(u(t_i)|u(t_{i-1}), u(t_{i+1}))$ and they can be found by Lemma 1.

3.1 Description of the algorithm

An iterative numerical algorithm to solve boundary value problem (1), (2) can be described as follows.

Initial step. A vector (which can be random or non-random)

$$U^{(0)} = (u^{(0)}(t_1), \ldots, u^{(0)}(t_N))$$

is taken as **initial**, where

$$u^{(0)}(t_1) = u(t_1) = u_*, \quad u^{(0)}(t_N) = u(t_N) = u^*.$$

For example, the points with the coordinates $(t_i, u^{(0)}(t_i))$ can be disposed on the straight line which connects the boundary points (t_1, u_*), (t_N, u^*).

Iterative step. We will consider two versions of the iterative step, to produce a random vector $U^{(n+1)}$ from the vector $U^{(n)}$.

(a) Random visiting scheme. First, according to a proposal distribution G on T_0 we choose a node $t \in T_0$. Then the value of $u^{(n+1)}(t)$ is simulated according to the distribution $P(u^{(n+1)}(t)|u^{(n)}(t-1), u^{(n)}(t+1))$. For all other values we set $u^{(n+1)}(s) = u^{(n)}(s)$, $t \neq s \in T_0$. The proposal distribution G is assumed to be strictly positive on T_0.

(b) Deterministic visiting scheme. The inner grid T_0 is scanned sequentially according to some fixed procedure (a sweep with deterministic visiting scheme) in such a way that all the nodes from T_0 must be, at least, once visited. For every visited node $t \in T_0$ the value of $u^{(n)}(t)$ is changed according to the conditional distribution $P(u^{(n)}(t)|u^{(n)}(t-1), u^{(n)}(t+1))$ (possible previous changes of $u^{(n)}(t-1)$, $u^{(n)}(t+1)$ during the current sweep should be taken into account). The configuration obtained at the end of the sweep will be taken as $U^{(n+1)}$.

After many iterative steps a configuration $U^{(n)}$ for large n is considered to be an approximation of U.

3.2 Convergence

The sequence $U^{(n)}$, $n = 1, 2, \ldots$, generated by the algorithm is a Markov chain with a continuous state space of $N * k$ dimension. Obviously, the distribution $P(U)$ is invariant for the Markov chain, and then (under some additional assumptions) the distributions of $U^{(n)}$ converge to $P(U)$.

The exact results can be obtained from the general ergodic theory for Markov chains presented, for example, in [5,6]. In particular, the following statements can be obtained like consequences of Orey's theorem.

Proposition 3.1 *Assume the matrices $R_{t_i}^{t_{i+1}}$, $i = 1, \ldots, N-1$, to be nonsingular. Then for any initial configuration $U^{(0)}$, $\|P(U^{(n)}) - P(U)\| \to 0$ as $n \to \infty$ (henceforth by $\|.\|$ we denote the total variation norm).*

Proposition 3.2 *Consider a boundary value problem (1), (2) for a time-invariant SDE with a stable matrix A, and assume the matrix R_α^β to be nonsingular. Then for an arbitrary initial configuration $U^{(0)}$, the algorithm on the basis of the Gibbs sampler converges to a solution of the boundary value problem, i.e., $\|P(U^{(n)}) - P(U)\| \to 0$ as $n \to \infty$.*

Remark 3.3 For a time-invariant SDE with a stable matrix A, the set $R_s^t(\mathbf{R}^k)$ does not depend on the values s and t for $s < t$. If the matrix R_α^β is singular, then existence of the solution is equivalent to the condition $x_\beta \in \exp((\beta - \alpha)A)x_\alpha + R_\alpha^\beta(\mathbf{R}^k)$, and the initial configuration cannot be arbitrarily chosen.

4 A general scheme for partial differential equations and nonlinear systems

The approach of Markov Chain Monte Carlo (MCMC) seems to be promising for the numerical solution of boundary value problems for stochastic differential equations of different types. There are several well-known schemes of MCMC like the Gibbs, the Metropolis and the Hastings methods. For stochastic differential equations (when there is a description of "local" dependency for the values of a simulated process or a field) the Gibbs sampler seems to be one of the most appropriate and natural.

Let us describe a general scheme. Consider a boundary value problem for a (partial nonlinear) SDE $Au(x) = w(x)$, where $x \in X \subset \mathbf{R}^d$ and $w(x)$ is white noise. After discretization, the boundary value problem can be reduced to another boundary value problem for a finite difference equation $\tilde{A}\tilde{u}(x_i) = \tilde{w}(x_i)$, where $\tilde{A}$ is a discrete approximation of the operator A on the grid $\tilde{X} \subset X$, $x_i \in \tilde{X}$, and $\tilde{w}(x_i)$ is a discrete white noise. The latter equation enables to find the conditional distributions $P(\tilde{u}(x_i)|\tilde{u}(x), x \in \partial(x_i))$ for $\tilde{u}(x_i)$ when the values of $\tilde{u}$ in the neighborhood $\partial(x_i) \subset \tilde{X}$ of the node x_i are given. Then the Gibbs sampler (like in Item 3.1) can be used for the approximate simulation of $\tilde{u}(x_i)$ on $\tilde{X}$: first, some initial configuration is fixed and then the nodes of the grid are scanned (according to a random or a deterministic visiting scheme) and the values in the nodes are updated according to the conditional distributions.

Remark 4.1

1. The following problems are significant for the considered algorithm: (a) the study of accuracy for discrete approximation, (b) an appropriate choice of the initial configuration and the stopping rule for the Gibbs sampler, (c) balancing the errors for discretization and the Gibbs sampler to diminish the total error of the result.

2. To optimize the algorithm, one can use such well-known techniques like the simultaneous updating for ensembles of nodes, synchronous and partially parallel algorithms, etc.

3. Conventional iteration methods for deterministic systems can be considered as an extreme case for the Gibbs samplers when white noise has zero intensity.

4. If the intensity of the white noise in the SDE tends to zero, then the result of simulation possibly converges to a solution for a certain problem of optimal control (cf. [2], p.153).

References

[1] S.M. Prigarin, Numerical solution of boundary value problems for linear systems of stochastic differential equations // Comp. Mathematics and Math. Physics.–1998.–V.38, N12.–P.1903-1908

[2] S.M. Prigarin, Spectral Models of Random Fields in Monte Carlo Methods. VSP, Utrecht, 2001

[3] S.M. Prigarin. On numerical solvability of boundary value problems for linear stochastic differential equations// Mathematical methods in stochastic simulation and experimental design. Proceedings of the 4th St.Petersburg Workshop on Simulation (2001, June 18 - June 22).- Saint Petersburg: Publishing House of Saint Petersburg University, 2001.- P.401-406

[4] G. Winkler, Image analysis, Random Fields and Dynamic Monte Carlo Methods. Springer, Berlin, 1995

[5] D. Revuz, Markov Chains. North-Holland Publishing Company, Amsterdam, 1975

[6] S.P. Meyn and R.L. Tweedie, Markov Chains and Stochastic Stability. Springer-Verlag, London, 1996