M5A44 COMPUTATIONAL STOCHASTIC PROCESSES

Professor G.A. Pavliotis

Department of Mathematics Imperial College London, UK
g.pavliotis@imperial.ac.uk

WINTER TERM 2016-2017
18/01/2015
• Lectures: Mondays, 12:00-14:00, Huxley 340, Wednesdays 10:00-11:00, Huxley 658.

• Office Hours: Wednesdays 14:00-16:00 or by appointment. Huxley 621.

• Course webpage: http://www.ma.imperial.ac.uk/~pavl/comp_stoch_proc.htm

• Text: Lecture notes, available from the course webpage. Also, recommended reading from various textbooks/review articles.
This is an introductory course on computational stochastic processes, aimed towards 4th year, MSc and MRes students in applied mathematics, statistics and theoretical physics.

Prior knowledge of basic stochastic processes in continuous time, scientific computing and a programming language such as matlab, R or C is assumed.
PART I: INTRODUCTION

- Random number generators.
- Random variables, probability distribution functions.
- Introduction to Monte Carlo techniques.
- Variance reduction methodologies for Monte Carlo simulations.

SIMULATION OF STOCHASTIC PROCESSES

- Introduction to stochastic processes.
- Brownian motion and related stochastic processes and their simulation.
- Gaussian stochastic processes.
- Karhunen-Loeve expansion and simulation of Gaussian random fields.
- Numerical solution of ODEs and PDEs with random coefficients.
3. NUMERICAL SOLUTION OF STOCHASTIC DIFFERENTIAL EQUATIONS.

- Basic properties of stochastic differential equations.
- Itô’s formula and stochastic Taylor expansions.
- Examples of numerical methods: The Euler-Maruyama and Milstein schemes.
- Theoretical issues: convergence, consistency, stability. Weak versus strong convergence.
- Implicit schemes and stiff SDEs.
4. ERGODIC DIFFUSION PROCESSES
   ▶ Deterministic methods: numerical solution of the Fokker-Planck equation.
   ▶ Numerical calculation of the invariant measure/steady state simulation.
   ▶ Calculation of expectation values, the Feynman-Kac formula.

5. MARKOV CHAIN MONTE CARLO (MCMC).
   ▶ The Metropolis-Hastings and MALA algorithms.
   ▶ Diffusion limits for MCMC algorithms.
   ▶ Peskun and Tierney orderings and Markov Chain comparison.
   ▶ Optimization of MCMC algorithms.
   ▶ Bias correction and variance reduction techniques.
6. STATISTICAL INFERENCE FOR DIFFUSION PROCESSES.
   ▶ Estimation of the diffusion coefficient (volatility).
   ▶ Maximum likelihood estimation of drift coefficients in SDEs.
   ▶ Nonparametric and Bayesian techniques.

7. APPLICATIONS
   ▶ Probabilistic methods for the numerical solution of partial differential equations.
   ▶ Exit time problems.
   ▶ Molecular dynamics/computational statistical mechanics.
   ▶ Molecular motors, stochastic resonance.
Prerequisites

- Basic knowledge of ODEs and PDEs.
- Familiarity with the theory of stochastic processes.
- Stochastic differential equations.
- Basic knowledge of functional analysis and stochastic analysis.
- Numerical methods and scientific computing.
- Familiarity with a programming language: Matlab, C, Fortran....
Assessment

- Based on 1 project (25%) and a final exam (75%).
- The project will be (mostly) computational.
- The exam will be theoretical (e.g. analysis of algorithms).

Feedback

- There will be no assessed coursework.
- I will hand out 2 or 3 problem sheets.
- You do not have to hand them in. If you do, then I will mark them and return them to you.
- There will be a few problems classes (1 hour per 2 weeks).
Lectures

- Mostly whiteboard plus slides.
- Computer experiments in Matlab.
Bibliography

- Lecture notes will be provided for all the material that we will cover in this course. They will be posted on the course webpage.
- Books that cover (parts of) the contents of this course are
why Computational Stochastic Processes?

- Many models in science, engineering and economics are probabilistic in nature and we have to deal with uncertainty.
  - Numerical weather prediction and climate modeling.
  - Computational statistical physics, molecular dynamics simulations.
  - Econometrics, Black-Scholes theory etc.
  - Reduction in complexity/number of degrees of freedom always leads to stochastic models.

- Many deterministic problems can be solved more efficiently using probabilistic techniques.
  - Calculation of integrals in high dimensional spaces.
  - Numerical solution of elliptic and parabolic PDEs in high dimensions using the Feynman-Kac formula.
Monte Carlo Methods

- We are interested in computing expectations of some function \( f \) with respect to the probability distribution \( \pi \), possibly known up to the normalization constant:

\[
I = \mathbb{E}_{X \sim \pi}[f(X)] := \int f(x)\pi(x)\,dx.
\] (1)

- Suppose that the state space is restricted to the unit square \( \Omega = [0, 1]^d \). We use numerical quadrature to approximate \( I \) to compute the integral directly.

- Using a regular mesh of \( K \times K \ldots K \) points, using the standard midpoint rule one can show that the error will be \( O(K^{-2}) \) (provided \( f \) is smooth).

- The number of evaluations of \( f \) and \( \pi \) will be \( M = N^d \). Thus, in terms of the computational effort, the error will be \( O\left(N^{-\frac{2}{d}}\right)\).

- As \( d \) increases, the number of evaluations must grow exponentially to maintain the same error.

- This problem is known as the **curse of dimensionality**.
Suppose we can generate a sequence of iid samples $x_1, x_2, \ldots$ of $\pi$.

From the law of large numbers we get

$$\frac{1}{N} \sum_{n=1}^{N} f(x_n) \to \mathbb{E}[f(X)] = \int f(x) \pi(x) \, dx, \quad \text{as } N \to \infty,$$

where $X \sim \pi$.

Thus, we can approximate the integral $I$ using the approximation

$$I_N = \frac{1}{N} \sum_{n=1}^{N} f(x_n),$$

knowing that, as $N \to \infty$, $I_N$ converges to $I$. 

As we will see, the rate of convergence is $O(N^{-1/2})$, which basically means that

$$\text{error} \sim \frac{1}{\sqrt{\text{computational cost}}}.$$  

This is much slower than that of deterministic methods, however it is (almost) independent of the number of dimensions.

Monte Carlo methods do not suffer from the curse of dimensionality.
Consider a system of (classical) particles (atoms, agents....).

The state of the system is described by their position and momenta: \( \mathbf{q}, \mathbf{p} = \{q_i, p_i : i = 1, \ldots k\}, q_i = (q_{i1}, q_{i2}, q_{i3}) \) (similarly for \( p_i \)).

The dynamics is described by the Hamiltonian function

\[
H(\mathbf{q}, \mathbf{p}) = \frac{1}{2} \mathbf{p}^T M^{-1} \mathbf{p} + U(\mathbf{q}).
\]

where \( M \) denotes the mass tensor.

The potential \( V(x) \) consists of a sum of pairwise interactions. A well known model for a classical fluid is that of the Lennard-Jones potential:

\[
U(x) = \sum_{i,j} \Phi(|x_i - x_j|), \quad \Phi(r) = 4\epsilon \left[ \left( \frac{\sigma}{r} \right)^{12} - \left( \frac{\sigma}{r} \right)^{6} \right].
\]
The equilibrium statistical mechanics of this system of interacting particles is described by the Boltzmann distribution:

$$\pi(x) = \frac{1}{Z} e^{-\beta U(x)},$$  \hspace{1cm} (2)

where $\beta = (k_B T)^{-1}$ is the inverse temperature and the normalization constant $Z$ is the partition function

$$Z = \int_{\mathbb{R}^3} e^{-\beta U(x)} \, dx.$$  \hspace{1cm} (3)

All the thermodynamic information about our system is given by the free energy:

$$F = -\beta^{-1} \ln Z.$$  \hspace{1cm} (4)

To calculate equilibrium quantities of interest we need to sample from the Boltzmann distribution (2).
We need to be able to calculate $Z(\beta)$, which is a very high dimensional integral.

We want to be able to calculate expectations with respect to $\pi(x)$:

$$\mathbb{E}_\pi f(x) = \int_{\mathbb{R}^3k} f(x) \pi(x) \, dx.$$ 

One method for doing this is by simulating a diffusion process (MCMC):

$$dX_t = -\nabla U(X_t) \, dt + \sqrt{2\beta^{-1}} \, dW_t.$$
Some interesting questions

- How many experiments should we do?
- How do we compute expectations with respect to stationary distributions?
- Can we exploit the problem structure to speed up the computation?
- How do we efficiently compute probabilities of rare events?
- How do we estimate the sensitivity of a stochastic model to changes in a parameter?
- How do we use stochastic simulation to optimize our choice of decision parameters?
Example: Brownian motion in a periodic potential

- Consider the Langevin dynamics in a smooth periodic potential

\[ \ddot{q} = -\nabla V(q) - \gamma \dot{q} + \sqrt{2\gamma \beta^{-1}} \dot{W}, \tag{5} \]

where \( W(t) \) denotes standard \( d \)-dimensional Brownian motion, \( \gamma > 0 \) the friction coefficient and \( \beta \) the inverse temperature (strength of the noise).

- Write as a first order system:

\[
\begin{align*}
 dq_t &= p_t \, dt, \\
 dp_t &= -\nabla q V(q_t) \, dt - \gamma p_t \, dt + \sqrt{2\gamma \beta^{-1}} \, dW_t. 
\end{align*}
\tag{6a} \tag{6b} \]
At long times the solution to $q_t$ performs an effective Brownian motion with diffusion coefficient $D$, $\mathbb{E} q_t^2 \sim 2Dt$, $t \ll 1$.

We compute the diffusion coefficient for $V(q) = \cos(q)$ using the formula

$$D = \lim_{t \to \infty} \frac{\text{Var}(q_t)}{2t},$$

as a function of $\gamma$, at a fixed temperature $\beta^{-1}$.

The SDE (6) has to be solved numerically using an appropriate scheme.

The numerical simulations suggest that

$$D \sim \frac{1}{\gamma},$$

for both $\gamma \ll 1$ and $\gamma \gg 1$. (7)
\[ \langle x^2 \rangle / 2t vs t \]

\[ D vs \gamma \]

Figure: Variance and effective diffusion coefficient for various values of $\gamma$. 
FIG. 1. \( \delta^{18}O \) record from the NGRIP ice core during the last glacial period.
• Fit this data to a bistable SDE

\[ dx = -V'(x; a) \, dt + \sigma \, dW, \quad V(x) = \sum_{j=1}^{4} a_j x^j. \]  

(8)

• Estimate the coefficients in the drift from the paleoclimatic data using the unscented Kalman filter.

• the resulting potential is highly asymmetric.
FIG. 8. Potential derived by least-squares fit from the probability density of the ice-core data (solid) together with probability densities of the model (dashed) and the data (dotted).
Fit this data to the generalized CIR (constant elasticity of variance) model

\[ dR_t = \alpha (\beta - R_t) \, dt + \sigma R_t^\gamma \, dW_t. \]  

(9)

Estimate the parameters \( \alpha, \beta, \sigma, \gamma. \)