Unbiased Simulation of Stochastic Differential Equations

Xiaolu Tan

Ceremade, University of Paris-Dauphine

Joint work with Pierre Henry-Labordère and Nizar Touzi

5th July, 2016

International Conference on Monte Carlo techniques @ Paris

Outline

1. Introduction
2. Unbiased estimators
   - A toy example
   - The unbiased estimators
3. Numerical examples

Outline

1. Introduction
2. Unbiased estimators
   - A toy example
   - The unbiased estimators
3. Numerical examples

Introduction

- Problem : Let X be a diffusion process defined by

$$X_t = x_0 + \int_0^t \mu(s, X_s)ds + \int_0^t \sigma(s, X_s)dW_s,$$

we aim to compute

$$\mathbb{E}\big[g(X_T)\big].$$

- Monte-Carlo methods
  - Time discretization method, MLMC, etc.
  - Exact simulation of X_T (Beskos, Roberts, Papaspiliopoulos, Joudain, Sbai, etc.)
  - Unbiased simulation method :
    - Randomized MLMC method (Rhee and Glynn, etc.).
    - Parametrix method (Bally and Kohatsu-Higa, Andersson and Kohatsu-Higa).

Outline

1. Introduction
2. Unbiased estimators
   - A toy example
   - The unbiased estimators
3. Numerical examples

A toy example

- Let X be defined by

$$X_t = x_0 + \int_0^t \mu(s, X_s)ds + W_t.$$

- The Feynmann-Kac formula : $\mathbb{E}[g(X_T)] = u(0, x_0)$, where u is solution of the linear parabolic PDE : $u(T, \cdot) = g(\cdot)$,

$$\partial_t u(t,x) + \mu(t,x) Du(t,x) + \frac{1}{2} D^2 u(t,x) = 0.$$

- Rewrite the PDE : for some fixed constant b

$$\partial_t u(t,x) + bDu(t,x) + \frac{1}{2} D^2 u(t,x) + \big(\mu(t,x) - b\big) Du(t,x) = 0.$$

A toy example

- We introduce $\widehat{X}_t = x_0 + bt + W_t$ to represent the solution of PDE

$$\partial_t u(t,x) + bDu(t,x) + \frac{1}{2}D^2 u(t,x) + \big(\mu(t,x) - b\big)Du(t,x) = 0,$$

- Using again Feynmann-Kac formula, it follows

$$\begin{aligned} u(0,x_0) &= \mathbb{E}\Big[g(\widehat{X}_T) + \int_0^T \big(\mu(s,\widehat{X}_s) - b\big)Du(s,\widehat{X}_s)ds\Big] \\ &= \mathbb{E}\Big[e^{\beta T} g(\widehat{X}_T) e^{-\beta T} + \int_0^T \big(\mu(s,\widehat{X}_s) - b\big)Du(s,\widehat{X}_s)\frac{e^{\beta s}}{\beta}\beta e^{-\beta s} ds\Big] \\ &= \mathbb{E}\Big[e^{\beta T} g(\widehat{X}_T)\mathbb{1}_{\{T_1 \geq T\}} \\ &\qquad + \big(\mu(T_1,\widehat{X}_{T_1}) - b\big)Du(T_1,\widehat{X}_{T_1})\frac{e^{\beta T_1}}{\beta}\mathbb{1}_{\{T_1 < T\}}\Big], \end{aligned}$$

where $T_1 := T \wedge \tau_1$ for $\tau_1 \sim \mathcal{E}(\beta)$ $(\tau_1 \perp W)$.

A toy example

- Recall the automatic differentiation formula :

$$\partial_x \mathbb{E}\big[\phi(x+bt+W_t)\big] \;=\; \mathbb{E}\Big[\phi(x+bt+W_t)\frac{W_t}{t}\Big],$$

it follows that, with $\Delta T_1 := T_1$,

$$Du(0,x_0) = \mathbb{E}\Big[e^{\beta \Delta T_1}\frac{\Delta W_1}{\Delta T_1}\Big(g(\widehat{X}_T)\mathbb{1}_{\{T_1\geq T\}} + \frac{\mu(T_1,\widehat{X}_{T_1})-b}{\beta}Du(T_1,\widehat{X}_{T_1})\mathbb{1}_{\{T_1<T\}}\Big)\Big].$$

- Denote $T_2 := (\tau_1+\tau_2)\wedge T$ and $\Delta T_2 := T_2 - T_1$, then

$$Du(T_1,\widehat{X}_{T_1}) = \mathbb{E}_{T_1}\Big[e^{\beta \Delta T_2}\frac{\Delta W_2}{\Delta T_2}\Big(g(\widehat{X}_T)\mathbb{1}_{\{T_2\geq T\}} + \frac{\mu(T_2,\widehat{X}_{T_2})-b}{\beta}Du(T_2,\widehat{X}_{T_2})\mathbb{1}_{\{T_2<T\}}\Big)\Big].$$

A toy example

- Let $(\tau_i)_{i\geq 1}$ be i.i.d. and $T_k := T \wedge \sum_{i=1}^{k} \tau_i$, denote $\Delta T_{k+1} := T_{k+1} - T_k$ and $N_T := \max\{k \ : T_k < T\}$, and

$$\psi_n \;\; := \;\; e^{\beta T_{n+1}} \Big(\prod_{k=1}^{N_T \wedge n} \frac{(\mu(T_k, \widehat{X}_{T_k}) - b)\Delta W_{k+1}}{\beta \Delta T_{k+1}} \Big) \Big(g(\widehat{X}_T) \mathbb{1}_{\{N_T \leq n\}} + \frac{\mu - b}{\beta} Du(T_{n+1}, \widehat{X}_{T_{n+1}}) \mathbb{1}_{\{N_T > n\}} \Big).$$

- Then

$$u(0, x_0) = \mathbb{E}[\psi_0] = \mathbb{E}[\psi_1] = \cdots = \mathbb{E}[\psi],$$

where

$$\psi \;\; := \;\; \lim_{n\to\infty} \psi_n \;\; = \;\; e^{\beta T} g(\widehat{X}_T) \prod_{k=1}^{N_T} \frac{(\mu(T_k, \widehat{X}_{T_k}) - b)\Delta W_{k+1}}{\beta \Delta T_{k+1}}.$$

A toy example

The constant b is arbitrary, one can choose it in an adaptive way, so that

$$\tilde{\psi} := e^{\beta T} g(\widehat{X}_T) \prod_{k=1}^{N_T} \frac{(\mu(T_k, \widehat{X}_{T_k}) - \mu(T_{k-1}, \widehat{X}_{T_{k-1}}))\Delta W_{k+1}}{\beta \Delta T_{k+1}},$$

where

$$\widehat{X}_{T_{k+1}} := \widehat{X}_{T_k} + \mu(T_k, \widehat{X}_{T_k})\Delta T_{k+1} + \Delta W_{k+1}.$$

And one has

$$\mathbb{E}\big[g(X_T)\big] = u(0, x_0) = \mathbb{E}[\tilde{\psi}].$$

General case

- Consider a general diffusion process

$$X_t = x_0 + \int_0^t \mu(s, X_s)ds + \int_0^t \sigma(s, X_s)dW_s,$$

then Feynmann-Kac formula leads to $\mathbb{E}[g(X_T)] = u(0, x_0)$:

$$\partial_t u(t,x) + \mu \cdot Du(t,x) + \frac{1}{2}\sigma\sigma^\top : D^2 u(t,x) = 0.$$

- The second order automatic differentiation formula

$$\partial^2_{xx}\mathbb{E}\big[\phi(x + bt + \sigma W_t)\big] = \mathbb{E}\Big[\phi\big(x + bt + \sigma W_t\big)(\sigma^\top)^{-1}\frac{W_t W_t^\top - t}{t^2}\sigma^{-1}\Big].$$

General case

Let

$$\widehat{X}_{T_{k+1}} := \widehat{X}_{T_k} + \mu(T_k, \widehat{X}_{T_k})\Delta T_{k+1} + \sigma(T_k, \widehat{X}_{T_k})\Delta W_{k+1}.$$

$$\psi := e^{\beta T} g(\widehat{X}_T) \prod_{k=1}^{N_T} (\mathcal{W}_k^1 + \mathcal{W}_k^2),$$

where

$$\mathcal{W}_k^1 := \frac{(\mu(T_k, \widehat{X}_{T_k}) - \mu(T_{k-1}, \widehat{X}_{T_{k-1}})) \cdot (\sigma_{T_k}^\top)^{-1} \Delta W_{k+1}}{\beta \Delta T_{k+1}},$$

$$\mathcal{W}_k^2 := (\sigma_{T_k}^\top)^{-1} \frac{(\sigma\sigma_{T_k}^\top - \sigma\sigma_{T_{k-1}}^\top) : (\Delta W_{k+1} \Delta W_{k+1}^\top - \Delta T_{k_1} I_d)}{2\beta \Delta T_{k+1}^2} \sigma_{T_k}^{-1}$$

It follows that

$$\mathbb{E}\big[g(X_T)\big] = u(0, x_0) = \mathbb{E}[\psi].$$

Extensions

- Other schemes for $\widehat{X}$ (rather than the Euler scheme). E.g. for $dX_t = \sigma(t, X_t)dW_t$,

$$d\widehat{X}_t = \big(\sigma(T_k, \widehat{X}_{T_k}) + \partial_x \sigma(T_k, \widehat{X}_{T_k})(\widehat{X}_t - \widehat{X}_{T_k})\big) dW_t.$$

- The path-dependent case : for some $0 = t_0 < \cdots < t_n = T$,

$$dX_t = \mu(t, X_{t_1 \wedge t}, \cdots, X_{t_n \wedge t})dt + \sigma_0 dW_t,$$

we aim to compute

$$\mathbb{E}\big[g(X_{t_1}, \cdots, X_{t_n})\big].$$

Extension and comparison

- Doumbia, Oudjane and Warin (2016), essentially change the exponential distribution to gamma distribution.
  - Comparing to Andersson and Kohatsu-Higa.

- Branching process, towards to the nonlinear PDE with polynomial nonlinearity :

$$\partial_t u + \mathcal{L}u + \sum_{\ell \in L} a_\ell u^{\ell_0} (Du)^{\ell_1} (D^2 u)^{\ell_2} = 0.$$

Outline

1 Introduction

2 Unbiased estimators
- A toy example
- The unbiased estimators

3 Numerical examples

A d-dimensional example (d=4)

Let $\sigma_0 \in S_4$ such that $\sigma_{ij} = 1_{i=j} + \frac{1}{2} 1_{i \neq j}$, $\quad X = (X^1, \cdots, X^4)$ be given by

$$dX_t \;=\; \mu(t, X_t)dt \;+\; \sigma_0 dW_t, \quad X_0^i = 0,\; i = 1, \cdots, 4,$$

where $\mu_i(t, x_1, \cdots, x_4) = 0.1\Big(\sqrt{\frac{3}{4}\exp(x_i) + \frac{1}{4}\overline{\exp(x)}} - 1\Big) - \frac{1}{8}$, and $\overline{\exp(x)} := (e^{x_1} + \cdots + e^{x_4})/4$.
Let $0 = t_0 < \cdots < t_{10} = 1$, we consider :

$$V_0 := \mathbb{E}\Big[\Big(\frac{1}{40}\sum_{k=1}^{10}\sum_{i=1}^{4} e^{X_{t_k}^i} - K\Big)_+\Big].$$

- Comparison with MLMC (Giles (2008)) methods.

A 4-dimensional example

	Mean value	Statistical error	Computation time
US ($N = 10^5$)	0.382186	0.00247547	0.769847
MLMC	0.381071	0.00167112	2.07589
US ($N = 10^6$)	0.382846	0.000762393	7.65796
MLMC	0.383107	0.000535905	10.8444
US ($N = 10^7$)	0.383282	0.000244861	85.0265
MLMC	0.383653	0.00017245	104.223

Table : US denotes our unbiased simulation algorithm with $\beta = 0.05$, the computation times are expressed in second.