J. Math. Anal. Appl. 458 (2018) 393–418

Contents lists available at ScienceDirect

Journal of Mathematical Analysis and Applications

www.elsevier.com/locate/jmaa

Variance reduction for discretised diffusions via regression ☆

Denis Belomestny [a,b], Stefan Häfner [c], Tigran Nagapetyan [d], Mikhail Urusov [a,*]

[a] University of Duisburg-Essen, Essen, Germany
[b] IITP RAS, Moscow, Russia
[c] PricewaterhouseCoopers GmbH, Frankfurt, Germany
[d] University of Oxford, UK

ARTICLE INFO

Article history:
Received 7 November 2016
Available online 14 September 2017
Submitted by U. Stadtmueller

Keywords:
Control variates
Monte Carlo methods
Regression methods
Stochastic differential equations
Weak schemes

ABSTRACT

In this paper we present a novel approach towards variance reduction for discretised diffusion processes. The proposed approach involves specially constructed control variates and allows for a significant reduction in the variance for the terminal functionals. In this way the complexity order of the standard Monte Carlo algorithm (ε^{-3} in the case of a first order scheme and $\varepsilon^{-2.5}$ in the case of a second order scheme) can be reduced down to $\varepsilon^{-2+\delta}$ for any $\delta \in [0, 0.25)$ with ε being the precision to be achieved. These theoretical results are illustrated by several numerical examples.

1. Introduction

Let $T > 0$ be a fixed time horizon. Consider a d-dimensional diffusion process $(X_t)_{t\in[0,T]}$ defined by the Itô stochastic differential equation

$$dX_t = \mu(X_t)\,dt + \sigma(X_t)\,dW_t, \quad X_0 = x_0 \in \mathbb{R}^d, \tag{1.1}$$

for Lipschitz continuous functions $\mu\colon \mathbb{R}^d \to \mathbb{R}^d$ and $\sigma\colon \mathbb{R}^d \to \mathbb{R}^{d\times m}$, where $(W_t)_{t\in[0,T]}$ is a standard m-dimensional Brownian motion. Recall that, since μ and σ are Lipschitz, the stochastic differential equation (1.1) has a strong solution, and pathwise uniqueness holds. Suppose we want to find a continuous function

$$u\colon [0,T]\times\mathbb{R}^d \to \mathbb{R},$$

☆ The work of Denis Belomestny is supported by the Russian Science Foundation project 14-50-00150. Stefan Häfner thanks the Faculty of Mathematics of the University of Duisburg-Essen, where this work was carried out. We thank the anonymous referee for the comments that helped improve the paper.

* Corresponding author.
E-mail addresses: denis.belomestny@uni-due.de (D. Belomestny), stefan.haefner@de.pwc.com (S. Häfner), nagapetyan@stats.ox.ac.uk (T. Nagapetyan), mikhail.urusov@uni-due.de (M. Urusov).

http://dx.doi.org/10.1016/j.jmaa.2017.09.002
0022-247X/

which has a continuous first derivative with respect to the time argument and continuous first and second derivatives with respect to the components of the space argument on $[0,T)\times\mathbb{R}^d$ such that it solves the partial differential equation

$$\frac{\partial u}{\partial t}+\mathcal{L}u=0 \quad \text{on } [0,T)\times\mathbb{R}^d, \tag{1.2}$$

$$u(T,x)=f(x) \quad \text{for } x\in\mathbb{R}^d, \tag{1.3}$$

where f is a given continuous function on $\mathbb{R}^d$. Here and in what follows, t denotes the time argument, x denotes the space argument of u, and $\mathcal{L}$ is the differential operator associated with the equation (1.1):

$$(\mathcal{L}u)(t,x):=\sum_{i=1}^{d}\mu^i(x)\frac{\partial u}{\partial x_i}(t,x)+\frac{1}{2}\sum_{i,j=1}^{d}(\sigma\sigma^\top)^{ij}(x)\frac{\partial^2 u}{\partial x_i\partial x_j}(t,x),$$

where $\sigma^\top$ denotes the transpose of σ, and the components of μ and $\sigma\sigma^\top$ (and later the ones of σ) are denoted by superscripts. Under appropriate conditions on μ, σ and f, there is a solution of the Cauchy problem (1.2)–(1.3), which is unique in the class of solutions satisfying certain growth conditions, and it has the following Feynman–Kac stochastic representation

$$u(t,x)=\mathbb{E}[f(X_T^{t,x})]$$

(see Section 5.7 in [4]), where $X^{t,x}$ denotes the solution started at time t in point x. Moreover it holds (see e.g. Newton [12])

$$\mathbb{E}[f(X_T^{0,x})|X_t^{0,x}]=u(t,X_t^{0,x}), \quad \text{a.s.}$$

for $t\in[0,T]$ and

$$f(X_T^{0,x})=\mathbb{E}[f(X_T^{0,x})]+M_T^*, \quad \text{a.s.} \tag{1.4}$$

with

$$M_T^*:=\int_0^T\nabla_x u(t,X_t^{0,x})\,\sigma(X_t^{0,x})\,dW_t=\int_0^T\sum_{i=1}^{d}\frac{\partial u}{\partial x_i}(t,X_t^{0,x})\sum_{j=1}^{m}\sigma^{ij}(X_t^{0,x})\,dW_t^j. \tag{1.5}$$

The standard Monte Carlo (SMC) approach for computing $u(0,x)$ at a fixed point $x\in\mathbb{R}^d$ consists of three steps. First an approximation $\overline{X}_T$ for $X_T^{0,x}$ is constructed via a time discretisation of the equation (1.1) (we refer to [5] for a nice overview of various discretisation schemes). Next N_0 independent copies of the approximation $\overline{X}_T$ are generated and finally a Monte Carlo estimate V_{N_0} is defined as an average of the values of f at simulated points:

$$V_{N_0}:=\frac{1}{N_0}\sum_{i=1}^{N_0}f\left(\overline{X}_T^{(i)}\right).$$

In the computation of $u(0,x)=\mathbb{E}[f(X_T^{0,x})]$ by the SMC approach there are two types of error inherent: a discretisation error $\mathbb{E}[f(X_T^{0,x})]-\mathbb{E}[f(\overline{X}_T)]$ and a Monte Carlo (statistical) error, which results from the substitution of $\mathbb{E}[f(\overline{X}_T)]$ with the sample average V_{N_0}. The aim of variance reduction methods is to reduce the statistical error. For example, in the so-called control variate variance reduction approach one looks