On randomized Euler scheme for SDEs with drift in integral form and its connection with SGD

Michał Sobieraj
AGH University of Krakow
`sobieraj@agh.edu.pl`

Coauthor(s): Paweł Przybyłowicz

In this presentation, we investigate strong approximation of solutions of the following stochastic differential equations

$$\begin{cases} \mathrm{d}X(t) = a(X(t))\mathrm{d}t + b(X(t))\mathrm{d}W(t), \ t \in [0,T], \\ X(0) = \eta, \end{cases} \tag{1}$$

where $d, m \in \mathbb{N}, \eta \in \mathbb{R}^d$, W is a m-dimensional Wiener process, $T \in [0, +\infty)$, and a is in the following integral form

$$a(x) = \int\limits_{\mathcal{T}} H(t,x)\mu(\mathrm{d}t),$$

such that $(\mathcal{T}, \mathcal{A}, \mu)$ is probability space (see [1]).

Hereafter, in a certain class of coefficients $H : \mathcal{T} \times \mathbb{R}^d \to \mathbb{R}^d$, $b : [0,T] \times \mathbb{R}^d \to \mathbb{R}^{d\times m}$, we investigate the upper L^p-error bound of introduced randomized Euler scheme $X^{RE}_{n,M}$ for pointwise approximation of solutions $X(T)$ of (1). Next, we discuss the connection between various variants of gradient descent algorithms (see [2]) and Euler schemes for the approximation of the solutions of differential equations. In particular, we focus on introduced randomized euler scheme $X^{RE}_{n,M}$ as a variant of perturbed stochastic gradient descent algorithm (see [3] where the case of fractional Wiener noise case was considered).

At the end we present results of numerical experiments performed on GPUs, where we provided a suitable implementation in CUDA C and Python to check correlation between estimated L^p-error and informational cost. Finally, the practical example of optimization problem which is solved with randomized Euler scheme is presented as well.

[1] S. N. Cohen, R. J. Elliott. (2015) *Stochastic Calculus and Applications, 2nd. ed.*, Springer.

[2] S. Ruder, *An overview of gradient descent optimization algorithms*, https://arxiv.org/abs/1609.04747

[3] A. Lucchi, F. Proske, A. Orvieto, F. Bach, H. Kersting, *On the Theoretical Properties of Noise Correlation in Stochastic Optimization*, https://arxiv.org/abs/2209.09162