Euler-Maruyama scheme for the SDE driven by stable process

Lihu Xu

University of Macau

August 28-September 1, 2023

The 2nd HKSIAM Biennial Meeting

Outline

- A general framework of stochastic approximation
- EM Scheme of SDE driven by stable process
- The optimal convergence rate
- Future work

Contents

Lindeberg method

Let $\xi_{n,1}, ..., \xi_{n,n}$ be a sequence of independent random variables such that

$$\mathbb{E}\xi_{n,k} = 0 \ \ \forall\ k, \quad \sum_{k=1}^{n} \mathbb{E}\xi_{n,k}^2 = 1.$$

Denote

$$S_n = \sum_{k=1}^{n} \xi_{n,k}.$$

Let $\eta_{n,1}, ..., \eta_{n,n}$ be a sequence of independent random variables such that $\eta_{n,k}$ is a normal random variable with $\mathbb{E}\eta_{n,k} = 0$ and $\mathbb{E}|\eta_{n,k}|^2 = \mathbb{E}|\xi_{n,k}|^2$.

Lindeberg method (ctd)

Denote

$$S_n^{(0)} = \xi_{n,1} + ... + \xi_{n,n}, \qquad S_n^{(1)} = \eta_{n,1} + \xi_{n,2} + ... + \xi_{n,n},$$

$$..., ...$$

$$S_n^{(n)} = \eta_{n,1} + ... + \eta_{n,n}.$$

We have

$$S_n^{(n)} \sim N(0, 1).$$

Lindeberg's method

Let $h \in C^3(\mathbb{R})$, we have

$$|\mathbb{E}h(S_n) - N(h)| = |\mathbb{E}h(S_n^{(0)}) - \mathbb{E}h(S_n^{(n)})| \leq \sum_{k=1}^{n} |\mathbb{E}[h(S_n^{(k)}) - h(S_n^{(k-1)})]|.$$

Denote

$$Y_{n,k} = \eta_{n,1} + ... + \eta_{n,k-1} + \xi_{n,k+1} + ... + \xi_{n,n},$$

then

$$S_n^{(k)} = Y_{n,k} + \eta_{n,k}, \quad S_n^{(k-1)} = Y_{n,k} + \xi_{n,k}.$$

Now we have

$$\mathbb{E}[h(S_n^{(k)}) - h(S_n^{(k-1)})] = \mathbb{E}[h(S_n^{(k)}) - h(Y_{n,k})] - \mathbb{E}[h(S_n^{(k-1)}) - h(Y_{n,k})].$$

Lindeberg's method (ctd)

If $|h'''(x)| \le C$ for all x, by third order Taylor expansion to $h(S_n^{(k)}) - h(Y_{n,k})$ and $h(S_n^{(k-1)}) - h(Y_{n,k})$, we have

$$\begin{aligned} |\mathbb{E}[h(S_n)] - N(h)| &\le \sum_{k=1}^{n} \left| \mathbb{E}[h(S_n^{(k)}) - h(S_n^{(k-1)})] \right| \\ &\le \frac{1}{6}\|h'''\| \left(\sum_{k=1}^{n} \mathbb{E}|\eta_{n,k}|^3 + \sum_{k=1}^{n} \mathbb{E}|\xi_{n,k}|^3 \right). \end{aligned}$$

When $X_1, ..., X_n$ be i.i.d. r.v. with $\mathbb{E}|X_i|^3 < \infty$, then $\xi_{n,k} = \frac{X_k}{\sqrt{n}}$ and we have

$$|\mathbb{E}[h(S_n)] - N(h)| \le C \frac{\|h'''\|}{\sqrt{n}}.$$

A new point of view

$$
\begin{aligned}
\mathbb{E}[h(S_n^{(k)}) - h(S_n^{(k-1)})] &= \mathbb{E}[h(S_n^{(k)}) - h(Y_{n,k})] - \mathbb{E}[h(S_n^{(k-1)}) - h(Y_{n,k})] \\
&= \mathbb{E}\{\mathbb{E}[h(Y_{n,k} + \eta_{n,k})|Y_{n,k}] - h(Y_{n,k})\} \\
&\quad - \mathbb{E}\{\mathbb{E}[h(Y_{n,k} + \xi_{n,k})|Y_{n,k}] - h(Y_{n,k})] \\
&= \mathbb{E}\{Ph(Y_{n,k}) - h(Y_{n,k})\} - \mathbb{E}\{Qh(Y_{n,k}) - h(Y_{n,k})\}
\end{aligned}
$$

where $Ph(x) = \mathbb{E}[h(x + \eta_{n,k})]$ and $Qh(x) = \mathbb{E}[h(x + \xi_{n,k})]$.

A universal approximation theorem: P. Chen, Q. M. Shao and X. (2022+)

Theorem 0 (General framework)

Let $N \geq 2$ be a natural number and let $h : E \to \mathbb{R}$ be a measurable function such that: (1). $\mathbb{E}|h(X_t^x)| < \infty$ and $\mathbb{E}|h(Y_k^y)| < \infty$ for all $x \in E, y \in E$, $t \leq N$ and $k \leq N$; (2). the function $u_k(x) := \mathbb{E}h(X_k^x)$ for $k \geq 1$ satisfies $\mathbb{E}|\mathcal{A}^X u_k(Y_j)| < \infty$ and $\mathbb{E}|\mathcal{A}^X u_k(X_t^{Y_j})| < \infty$ for all $1 \leq j, k \leq N$ and $0 \leq t \leq 1$. Then

$$\mathbb{E}h(X_N) - \mathbb{E}h(Y_N) \;=\; \mathcal{I} + \mathcal{II} + \mathcal{III}, \tag{1}$$

A universal approximation theorem

Theorem 0 (General framework (ctd))

where $\mathcal{A}^X$ and $\mathcal{A}^Y$ are the infinitesimal generators of $(X_t)_{t\geq 0}$ and $(Y_k)_{k\geq 0}$ respectively, and

$$\mathcal{I} = \sum_{j=1}^{N-1} \mathbb{E}\big[\mathcal{A}^X u_{N-j}(Y_{j-1}) - \mathcal{A}^Y u_{N-j}(Y_{j-1})\big],$$

$$\mathcal{II} = \sum_{j=1}^{N-1} \mathbb{E}\int_0^1 \big[\mathcal{A}^X u_{N-j}(X_s^{Y_{j-1}}) - \mathcal{A}^X u_{N-j}(Y_{j-1})\big]\mathrm{ds},$$

$$\mathcal{III} = \mathbb{E}\big[h\big(X_1^{Y_{N-1}}\big) - h(Y_{N-1})\big] + \mathbb{E}\big[h(Y_N) - h(Y_{N-1})\big].$$

A remark about the theorem

To use the theorem, we need to

- choose the function family of h, e.g.
  - bounded measurable: TV metric
  - Lipschitz: Wasserstein-1 metric
- bound the three terms $\mathcal{I}$–$\mathcal{III}$: PDE method, heat kernel, Malliavin calculus.
- We have applied this framework to study the following problems: normal approximation, stable approximation, SGD approximation, SVRG approximation, EM scheme approximation,...

Contents

stochastic differential equation driven by stable noise

The SDE

$$\mathrm{dX_t = b(X_t)\,dt + dZ_t, \quad X_0 = x,}$$

where

- $x \in \mathbb{R}^d$ is the starting point,
- $(Z_t)_{t\geq 0}$ is a d-dimensional, rotationally invariant α-stable Lévy process with index $\alpha \in (1,2)$,
- b is Lischitz, there exist some $c > 0, K > 0$ such that for all x, y

$$\langle b(x) - b(y), x - y \rangle \leq -c|x-y|^2 + K.$$

The EM scheme

EM scheme:

$$Y_0 = x, \quad Y_{k+1} = Y_k + \eta b(Y_k) + \frac{\eta^{1/\alpha}}{\sigma} \widetilde{Z}_{k+1}, \quad k = 0, 1, 2, \ldots,$$

where

- $\widetilde{Z}_1, \widetilde{Z}_2, \cdots$ is an iid sequence with Pareto distribution, i.e.

$$\widetilde{Z}_1 \sim p(z) = \frac{c}{|z|^{\alpha+d}} 1_{(1,\infty)}(|z|),$$

- η is the step size.

Ergodicity of EM scheme and SDE

Under the condition: b is Lischitz, there exist some $c > 0, K > 0$ such that for all x, y

$$\langle b(x) - b(y), x - y \rangle \leq -c|x - y|^2 + K,$$

we have

- $(X_t)_{t \geq 0}$ is ergodic, denote the ergodic measure by μ,
- $(Y_k)_{k \geq 0}$ is ergodic, denote the ergodic measure by μ_η.

stochastic differential equation driven by stable noise

Theorem

(Chen, Deng, Schilling, X.) As b satisfies the above condition, there exists a constant C such that the following two statements hold:

1. *For every $N \geq 2$, one has*

$$W_1\big(\text{law}(X_{\eta N}), \text{law}(Y_N)\big) \leq C(1+|x|)\eta^{2/\alpha-1}.$$

2. *One has*

$$W_1\big(\mu, \mu_\eta\big) \leq C\eta^{2/\alpha-1},$$

where μ and μ_η are ergodic measures of $(X_t)_{t\geq 0}$ and $(Y_k)_{k\geq 0}$.

Contents

Some recent work on stable type processes

- Stable random fields: Xiao, Peligrad, Sang, Yang,...,...,
- Stable type processes: Chen, Kyprianou, Wang, Schilling, Song, Xiao, Yang, Zheng...,...,
- SDEs driven by stable processes: Deng, Kyprianou, Schilling, Wang, Zhai, Zhang, Zhang,...,...,
- EM scheme: Bao, Huang, Schilling, Yuan,...,...,

Contents

The strategy of the proof

- For a Lipschitz function h, define

$$P_t h(x) = \mathbb{E} h(X_t^x), \qquad Q_k h(x) = \mathbb{E} h(Y_k^x).$$

- The framework can be simplified in this special case as

$$P_{N\eta} h(x) - Q_N h(x) = \sum_{i=1}^{N} Q_{i-1} \big(P_\eta - Q_1\big) P_{(N-i)\eta} h(x). \tag{2}$$

- Need to estimate $\big(P_\eta - Q_1\big) P_t h(x)$:
  - the regularity of $P_t h(x)$ plays a crucial role,
  - we use Malliavin calculus to study it.

Subordination

- $Z_t := W_{S_t}$, where W_t is a d dimensional standard Brownian motion, $\{S_t\}_{t\geq 0}$ be an independent $\frac{\alpha}{2}$-stable subordinator.
- The equation can be rewritten as

$$\mathrm{d}X_t = b(X_t)\,\mathrm{d}t + \mathrm{d}W_{S_t}, \quad X_0 = x. \tag{3}$$

Given a sample path l_t from the subordinator S_t, consider the SDE:

$$\mathrm{d}X_t^l = b(X_t^l)\,\mathrm{d}t + \mathrm{d}W_{l_t}, \quad X_0 = x. \tag{4}$$

- $P_t h(x) = \mathbb{E}[h(X_t^x)] = \mathbb{E}[\mathbb{E}[h(X_t^{l,x})]|_{l=S}] = \mathbb{E}[P_t^l h(x)|_{l=S}]$.
- How to get the regularity of $P_t^l h(x)$?

A subordinator path and its approximation

- Given a sample path l_t from the subordinator S_t, it is cadlag and nondecreasing.
- For every $\epsilon \in (0,1)$, define its approximation as

$$l_t^\epsilon := \frac{1}{\epsilon}\int_t^{t+\epsilon} l_s \, \mathrm{d}s + \epsilon \mathrm{t}.$$

- Let γ_t^ϵ be the inverse function of l_t^ϵ, then

$$l^\epsilon_{\gamma_t^\epsilon} = t, \quad t \geq l_0^\epsilon \quad \text{and} \quad \gamma^\epsilon_{l_t^\epsilon} = t, \quad t \geq 0.$$

Time change (Zhang, SPA, 2013)

- By definition, γ_t^ϵ is absolutely continuous on $[l_0^\epsilon, \infty)$. Let us now define

$$Y_t^{l^\epsilon} := X_{\gamma_t^\varepsilon}^{l^\epsilon}, \quad t \geq l_0^\epsilon.$$

- Changing variables in (4) we see that for $t \geq l_0^\epsilon$,

$$Y_t^{l^\epsilon} = x + \int_{l_0^\epsilon}^t b\left(Y_s^{l^\epsilon}\right) \dot{\gamma}_s^\epsilon \, \mathrm{ds} + \mathrm{W_t} \tag{5}$$

($\dot{\gamma}_s^\epsilon$ denotes the derivative in s).

Time change and Malliavin calculus (Zhang, SPA, 2013)

- We apply Malliavin calculus to the SDE w.r.t. $Y_t^{l^\epsilon}$ and obtain the regularity of the associated semigroup.
- Recall

$$Y_t^{l^\epsilon} := X_{\gamma_t^\varepsilon}^{l^\epsilon}, \quad t \geq l_0^\epsilon,$$

transfer the regularity w.r.t. $Y_t^{l^\epsilon}$ to that w.r.t. $X_{\gamma_t^\varepsilon}^{l^\epsilon}$.

- Pass to the limit of $X_{\gamma_t^\varepsilon}^{l^\epsilon}$ to X_t^l as $\epsilon \to 0$, and obtain the regularity of $P_t^l h(x)$.

The rate $\eta^{2/\alpha-1}$ is optimal

We shall use OU stable process to verify that our convergence rate is optimal:

$$dX_t = -X_t dt + dZ_t.$$

Choose the Lipschitz function $h(x) = \frac{1}{M}\left(\frac{\sin x}{x} 1_{\{x\neq 0\}} + 1_{\{x=0\}}\right)$,

$$\begin{aligned}
& W_1(\mu, \mu_\eta) \\
\geq & \left|\mathbb{E}\left[h(Y_\eta)\right] - \mathbb{E}\left[h(\alpha^{-1/\alpha} Z_1)\right]\right| \\
= & \left|\int_{\mathbb{R}} \left(\frac{1}{2M}\int_{-1}^{1} e^{i\xi x}\,\mathrm{d}\xi\right)\mathbb{P}(Y_\eta \in \mathrm{dx}) - \int_{\mathbb{R}}\left(\frac{1}{2\mathrm{M}}\int_{-1}^{1} \mathrm{e}^{\mathrm{i}\xi\mathrm{x}}\,\mathrm{d}\xi\right)\mathbb{P}(\alpha^{-\frac{1}{\alpha}}\mathrm{Z}_1 \in \mathrm{d}
\right. \\
= & \left|\frac{1}{2M}\int_{-1}^{1}\mathbb{E}\left[e^{i\xi Y_\eta}\right]\mathrm{d}\xi - \frac{1}{2\mathrm{M}}\int_{-1}^{1}\mathbb{E}\left[\mathrm{e}^{\mathrm{i}\xi\alpha^{-1/\alpha}\mathrm{Z}_1}\right]\mathrm{d}\xi\right| \\
= & \frac{1}{2M}\left|\int_{-1}^{1}\left(\mathbb{E}\left[e^{i\xi Y_\eta}\right] - \mathbb{E}\left[e^{i\xi\alpha^{-1/\alpha}Z_1}\right]\right)\mathrm{d}\xi\right| \geq \Omega(\eta^{2/\alpha-1}).
\end{aligned}$$

Contents

Summary and the future work

Summary

- We introduce a probability approximation framework by modifying Lindeberg principle.
- We show by this framework that the EM scheme of SDE driven by stable process can approximate the ergodic measure of the SDE.

The future work

- The noise is multiplicative, we need a non-adaptive Malliavin calculus (Chen, X., Zhang and Zhang).
- The step size is decreasing (Chen, Jin, Xiao, X.).

P. Chen*, Q.M. Shao, L. Xu: A probability approximation framework: Markov process approach, Annals of Applied Probability, (2023).

P. Chen*, J. Lu*, L. Xu: Approximation to stochastic variance reduced gradient Langevin dynamics by stochastic delay differential equations, Applied Mathematics and Optimizations, (2022).

P. Chen*, C. Deng, R. Schilling, L. Xu: Approximation of the invariant measure of stable SDEs by an Euler–Maruyama scheme, Stochastic Processes and Their Applications.

X. Jin*, G. Pang, L. Xu, X. Xu*: An approximation to steady-state of M/Ph/n+M queue, Mathematics of Operations Research (minor revision), arXiv:2109.03623.

Thanks A Lot!