Multilevel-Langevin pathwise average for average for Gibbs approximation

Journées MAS 2022

M. EGÉA [1]

August 29, 2022

[1] LAREMA, Université d'Angers, E-mail: maxime.egea@univ-angers.fr

Introduction : Langevin diffusion and Gibbs measure

Let $U : \mathbb{R}^d \to \mathbb{R}$ be a differentiable function and consider the **SDE** :

$$\mathrm{d}X_t = -\nabla U(X_t)\mathrm{d}t + \sqrt{2}\mathrm{d}B_t,$$

where $(B_t)_{t\geq 0}$ is a d-dimensional Brownian motion.

Theorem

Under appropriate assumptions, the solution $(X_t)_{t\geq 0}$ has a unique invariant distribution π with density

$$\pi(x) = \frac{1}{Z_U}\exp(-U(x)),$$

π is called a Gibbs measure.

Objective

The goal is to estimate $\pi(f) := \int_{\mathbb{R}^d} f(x)\pi(\mathrm{d}x)$.

Standard Monte Carlo method: *Sample $X_1, X_2, \ldots, X_N$ i.i.d with distribution π and define*

$$\mathcal{M}(f) := \frac{1}{N}\sum_{i=1}^{N} f\left(X_i\right).$$

MSE:

$$\|\mathcal{M}(f) - \pi(f)\|_2^2 \leq \frac{\mathrm{Var}(f(X_1))}{N}.$$

Then, $\mathrm{Var}(f(X_1))\varepsilon^{-2}$ ($\approx d\varepsilon^{-2}$ if f is 1-Lipschitz and $\mathrm{Var}(X_1) \leq d$) samples are needed for an ε-approximation.

Multilevel method

Consider a family of random variables $(X_j)_{j\in\mathbb{N}}$ approximating X for J the number of "levels", observe that

$$\mathbb{E}[X_J] = \underbrace{\mathbb{E}[X_0]}_{coarse} + \sum_{j=1}^{J} \underbrace{\mathbb{E}[X_j - X_{j-1}]}_{correcting\ layer}.$$

Usually, $X_j := \bar{X}_N^{\gamma_r}$ is an Euler discretization of the **(SDE)** :

$$X_{n+1}^{\gamma_r} = x - \gamma_r \nabla U\left(X_n^{\gamma_r}\right) + \sqrt{2\gamma_r} Z_n,$$

where $(Z_n)_{n\in\mathbb{N}}$ is an $i.i.d$ sequence of d-dimensional standard Gaussian random variables.

Literature

Finite horizon approximation ($\mathbb{E}[f(X_T)]$)

- Giles - 2008 (Multilevel Monte Carlo),
- Lemaire, Pagès - 2014 (Romberg, Multilevel Monte Carlo).

Approximation of the invariant distribution $\pi(f)$

- Pagès, Panloup - 2018 (Ergodic Multilevel, Romberg, decreasing time step).
- Giles, et al. - 2020 (Multilevel Monte carlo),

Ergodic Multilevel : An occupation measure approach

We develop a multilevel procedure using the following property

$$\frac{1}{T}\int_0^T \delta_{X_s}\mathrm{d}s \underset{T\to+\infty}{\longrightarrow} \pi \qquad a.s.$$

Denote by

$$\nu_{\tau,T}^{\gamma}(f) := \frac{1}{T-\tau}\int_{\tau}^{T} f\left(\overline{X}_s^{\gamma}\right)\mathrm{d}s,$$

Ergodic Multilevel : the method

Definition

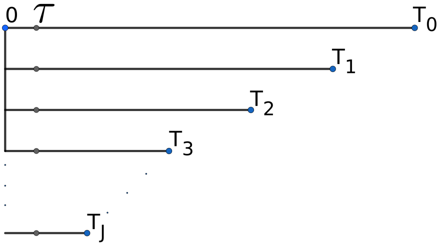
For $J \in \mathbb{N}$ the number of levels, $(T_j)_j$ a decreasing sequence of final time, $(\gamma_j)_j$ a decreasing sequence of time steps. Define the multilevel procedure

$$\mathcal{Y}(f) := \nu^{\gamma_0}_{\tau,T_0}(f) + \sum_{j=1}^{J} \nu^{\gamma_j}_{\tau,T_j}(f) - \nu^{\gamma_{j-1}}_{\tau,T_j}(f),$$

where Euler schemes are **coupled** in each level.

For example,
$\forall j \in \{1, \dots, J\}$:

- $\gamma_j = \gamma_0 2^{-j}$
- $T_j = T_0 2^{-j}$



Ergodic Multilevel : Mean squared error

Bias-variance decomposition :

$$
\begin{aligned}
\|\mathcal{Y}(f) - \pi(f)\|_2^2 &= \mathbb{E}[\mathcal{Y}(f) - \pi(f)]^2 + \mathrm{Var}(\mathcal{Y}(f)) \\
&\leq 2\underbrace{|\pi^{\gamma_J}(f) - \pi(f)|^2}_{\text{discretization error}} + 2\underbrace{\mathbb{E}[|\mathcal{Y}(f) - \pi^{\gamma_J}(f)|]^2}_{\text{longtime error}} + \mathrm{Var}(\mathcal{Y}(f)),
\end{aligned}
$$

Hypothesis

H_1 **(Convergence to equilibrium)** : There exist $\alpha > 0$ and $c_1 > 0$ such that for all $t \geq 0$,

$$|\mathbb{E}[f(X_t)] - \pi(f)| \leq c_1 e^{-\alpha t}. \tag{1}$$

H_2 **(weak error)** : There exists $c_2 > 0$ such that for all $\gamma \in (0, \eta_0]$,

$$|\pi(f) - \pi^\gamma(f)| \leq c_2 \gamma^\delta, \tag{2}$$

where δ is a positive number.

H_3 **(L^2 confluence)** : There exist $\rho \geq 1$ and a real constant c_3 such that for every $t \geq 0$,

$$\sup_{t\geq 0} \left\| \overline{X}_t^{\gamma/2} - \overline{X}_t^{\gamma} \right\|_2 \leq c_3 \gamma^{\frac{\rho}{2}}. \tag{3}$$

H_4 **(control of the moments)** : There exists a constant $c_4 > 0$ such that for all $\gamma \in (0, \eta_0]$,

$$\sup_{t>0} \left\| \overline{X}_t^{\gamma} \right\|_2 \leq c_4. \tag{4}$$

Ergodic Multilevel

Theorem

Assume H_1, H_2, H_3 and H_4 and ∇U is L-Lipschitz continuous. Let f be a Lipschitz continuous function. Then, for any $\varepsilon > 0$, there exist parameters such that

$$\|\mathcal{Y}(f) - \pi(f)\|_2 \leq C\varepsilon,$$

where C is a constant depending on c_1, c_2, c_3 and c_4. Furthermore, the complexity is

$$Compl(\mathcal{Y}) \simeq \begin{cases} \varepsilon^{-2}\log(\varepsilon^{-1})^3 & \text{if } \rho = 1, \\ \varepsilon^{-2} & \text{if } \rho > 1. \end{cases} \tag{5}$$

Ergodic Multilevel : Complexity

For example :

$$\tau = \frac{1}{\alpha}|\log(\varepsilon)|, \qquad J = \left\lfloor \frac{|\log(\varepsilon)|}{\log(2)} \right\rfloor, \qquad \gamma_j = \gamma_0 2^{-j},$$

$$T_j = \begin{cases} 2^{-j}\varepsilon^{-2}|\log(\varepsilon)|^2 & \text{if } \rho = 1, \\ 2^{-j\frac{\rho+1}{2}}\varepsilon^{-2} & \text{if } \rho > 1, \end{cases}$$

The complexity cost of the method is given by

$$Compl(\mathcal{Y}) = \frac{T_0}{\gamma_0} + \sum_{j=1}^{J} \frac{T_j}{\gamma_j} + \frac{T_j}{\gamma_{j-1}}.$$

What is the dependence on the dimension of the complexity ?

Application to uniformly convex setting

C_s : There exists $\alpha > 0$ such that for all $x, y \in \mathbb{R}^d$,

$$\langle \nabla U(x) - \nabla U(y), y - x \rangle \leq -\alpha |x - y|^2.$$

Proposition

Assume C_s and ∇U is L-Lipschitz continuous. Let $\gamma_0 \in (0, \frac{\alpha}{2L^2}]$. Then H_1, H_2, H_3 and H_4 hold with $\rho = 1$,

$$c_1(x) \leq C(|x| + \sqrt{d}) \qquad \text{and,} \qquad c_i \leq C\sqrt{d} \quad \text{for} \quad i \in \{2, 3, 4\},$$

where C does only depend on L and α.

Application to uniformly convex setting

Corollary

Assume C_s, ∇U is L-Lipschitz continuous and f is 1-Lipschitz continuous. Then for any $\varepsilon > 0$ there exist parameters such that $\mathcal{Y}(f)$ is an ε-approximation of $\pi(f)$ and

$$Compl(\mathcal{Y}) \leq d\varepsilon^{-2}\frac{CL^2}{\alpha^3},$$

where C is a universal constant.

Merci pour votre attention !

M. Egéa and F. Panloup
Multilevel-Langevin pathwise average for Gibbs approximation,
Hal preprint: hal-03345717, (2021).

References

A. Durmus and É. Moulines,
High-dimensional Bayesian inference via the unadjusted Langevin algorithm,
Bernoulli, volume 25 p.2854-2882, (2019).

W. Mou, N. Flammarion, M.J. Wainwright and P.L. Bartlett,
Improved bounds for discretization of Langevin diffusions: Near-optimal rates without convexity,
arXiv preprint:1907.11331,(2019).

S. Gadat, F. Panloup and C. Pellegrini
On the cost of Bayesian posterior mean strategy for log-concave models,
arXiv preprint:2010.06420, (2020).

References

M.B Giles
Multilevel monte carlo path simulation,
Operations research, volume 56 n°3 p.607-617, (2008).

M.B Giles and L. Szpruch
Antithetic multilevel Monte Carlo estimation for multi-dimensional SDEs without Lévy area simulation,
The Annals of Applied Probability, volume 24 n°4 p.1585-1620 (2014).

References

M.B Giles, L. Szpruch, M. Majka, M. S. Vollmer and K.C. Zygalakis
Multi-level Monte Carlo methods for the approximation of invariant measures of stochastic differential equations,
Statistics and Computing, volume 30 n°3 p.507-524 (2020).

G. Pagès and F. Panloup
Weighted multilevel Langevin simulation of invariant measures,
The Annals of Applied Probability, volume 28 n°6 p. 3358-3417 (2018).