



北京大学
PEKING UNIVERSITY

复杂随机过程和数值方法的遍历性和长时间误差

答辩人： 虚空若叶睦

院系： 数学科学学院

2025年5月26日

目录

- ① 采样问题和采样算法简介
- ② 路径积分分子动力学
- ③ 随机分组方法
- ④ 总结和致谢



Section 1

采样问题和采样算法简介



计算数学中的采样问题

- 计算数学中的采样 (sampling) 问题的出现源于对高维分布的近似.
- 给定目标分布 $\pi(x) \propto e^{-V(x)}$, 其中 $V(x)$ 是 \mathbb{R}^d 上的势能函数. 采样的目标是高效生成一系列样本点 $(X_n)_{n=0}^{\infty}$, 使得当 n 充分大时, X_n 的分布 $\mu_n \in \mathcal{P}(\mathbb{R}^d)$ 能快速且准确地收敛到目标分布 $\pi(x)$.
- 采样问题广泛地出现在计算物理, 数据科学和机器学习等领域. 典型的应用场景包括: 自由能的计算, 贝叶斯推断和扩散生成式模型.



采样算法简介

- 解决采样问题的核心是设计合适的采样算法. 在大部分采样算法中, 算法在每个迭代点 $X_n \in \mathbb{R}^d$ 附近收集 $V(x)$ 或 $\nabla V(x)$ 的信息, 并用于确定下一个迭代点 X_{n+1} 的位置.
- 采样算法的复杂度是指调取 $V(x)$ 和 $\nabla V(x)$ 的次数.
- 常用的采样算法包括:
 - Metropolis–Hastings 算法
 - Hamiltonian Monte Carlo (HMC)
 - 基于 Langevin 动力学的离散格式: Euler–Maruyama, MALA, randomized midpoint, BAOAB & UBU
 - 随机梯度采样算法: SG-LD, SG-HMC, RBM



采样算法的误差刻画

- 在本报告中, 我们主要研究基于 Langevin 动力学的采样算法. 遍历性是刻画 Langevin 动力学收敛速度的核心概念.
- 设随机过程 $(x_t)_{t \geq 0}$ 的对偶半群为 $(\mathcal{P}_t)_{t \geq 0}$, 即若 $x_0 \sim \nu$, 则 $x_t \sim \nu \mathcal{P}_t$. 若 π 是 $(x_t)_{t \geq 0}$ 的不变分布, 即

$$\pi \mathcal{P}_t = \mathcal{P}_t, \quad \forall t \geq 0,$$

且存在常数 $C, \lambda > 0$ 使得以及一个概率度量 $d(\cdot, \cdot)$, 使得

$$d(\nu \mathcal{P}_t, \pi) \leq C e^{-\lambda t} d(\nu, \pi), \quad \forall t \geq 0,$$

则称 $(x_t)_{t \geq 0}$ 具有**遍历性**.

- 常用的概率度量包括: 全变差, Wasserstein 距离和 KL 散度 (相对熵).



采样算法的误差刻画

- 为 Langevin 动力学 $(x_t)_{t \geq 0}$ 的离散格式指定步长 $h > 0$, 可以得到离散解 $(X_n)_{n=0}^\infty$. 采样误差与 Langevin 动力学的收敛速度和步长均相关.
- 例: \mathbb{R}^d 上的过阻尼 Langevin 动力学形如

$$dx_t = -\nabla U(x_t) + \sqrt{2}dB_t, \quad t \geq 0,$$

其中 $(B_t)_{t \geq 0}$ 为 Brown 运动. 它的 Euler-Maruyama 格式为

$$X_{n+1} = X_n - h\nabla U(X_n) + \sqrt{2h}(B_{(n+1)h} - B_{nh}), \quad n \geq 0.$$

- 上述离散解 $\{X_n\}_{n=0}^\infty$ 的误差可控制为: (定理 5, [BDMS18])

$$\mathcal{W}_2(\nu\tilde{P}_n^h, \pi) \leq Ce^{-\lambda nh} + Ch,$$

其中 C, λ 为常数, \tilde{P}_n^h 为 $(X_n)_{n=0}^\infty$ 的对偶半群.



采样算法的误差刻画

- 一般地, 假设 Langevin 动力学 $(x_t)_{t \geq 0}$ 的离散解为 $(X_n)_{n=0}^\infty$, 且其对偶半群为 \tilde{P}_n^h . 若存在一个满足 $\lim_{h \rightarrow 0} \varepsilon(h) = 0$ 的误差函数 $\varepsilon(h)$, 使得

$$d(\nu \tilde{P}_n^h, \pi) \leq C e^{-\lambda n h} + C \varepsilon(h),$$

则称 $(X_n)_{n=0}^\infty$ 的采样误差具有指数衰减性.

- 本报告的主题对于特定的高维 Langevin 动力学及其离散格式, 证明相应的遍历性和采样误差的指数衰减性.



博士期间的主要工作

主要研究了路径积分分子动力学 (path integral molecular dynamics) 和随机分组方法 (random batch method) 的遍历性和采样误差:

- ① Efficient sampling of thermal averages of **interacting quantum particle systems** with **random batches** (J. Chem. Phys., 2021)
- ② Ergodicity and long-time behavior of the **random batch method** for interacting particle systems (M3AS, 2023)
- ③ Error analysis of **time-discrete random batch method** for interacting particle systems and associated mean-field limits (IMA J. Numer. Anal., 2023)
- ④ Dimension-free ergodicity of **path integral molecular dynamics** (accepted by CiCP, 2023)



Section 2

路径积分分子动力学



正则系综平均的路径积分表示

- 路径积分分子动力学 (**path integral molecular dynamics**) 是计算量子系统的**正则系综 (canonical ensemble)** 平均的一套标准方法。
- \mathbb{R}^d 中的量子系统可由 Hamilton 算子

$$\hat{H} = -\frac{\Delta}{2} + V(\hat{x}),$$

来描述, 该系统在温度 T 下的正则系综由密度算子

$$\hat{\rho}_\beta = \frac{1}{\mathcal{Z}_\beta} e^{-\beta \hat{H}},$$

其中 $\beta = 1/(k_B T)$, 而 $\mathcal{Z}_\beta = \text{Tr} [e^{-\beta \hat{H}}]$ 是配分函数 (partition function).



正则系综平均的路径积分表示

- 计算 $\langle x | e^{-\beta \hat{H}} | x \rangle$ 本质上需要求解 \hat{H} 对应的抛物方程, 若直接使用空间离散或谱方法, 计算复杂度将随 d **指数级增长** (维数灾难).
- 路径积分提供了近似 $\langle x | e^{-\beta \hat{H}} | x \rangle$ 的一种方法, 且复杂度随 d **线性增长**.
- 给定 $x_1 \in \mathbb{R}^d$, 利用路径积分中的**时间切片法**, $\langle x_1 | e^{-\beta \hat{H}} | x_1 \rangle$ 可近似为

$$\begin{aligned} & \langle x_1 | e^{-\beta \hat{H}} | x_1 \rangle \\ & \approx \int \prod_{j=1}^N \langle x_j | e^{-\frac{\beta_N}{2} V} e^{\frac{\beta_N}{2} \Delta} e^{-\frac{\beta_N}{2} V} | x_{j+1} \rangle dx_2 \cdots dx_N \quad (\text{使用 Strang 分裂}) \\ & = \frac{1}{(2\pi\beta_N)^{\frac{dN}{2}}} \int \exp \left(-\frac{1}{2\beta_N} \sum_{j=1}^N |x_j - x_{j+1}|^2 - \beta_N \sum_{j=1}^N V(x_j) - \mathcal{E}_N(\mathbf{x}) \right) dx_2 \cdots dx_N, \end{aligned}$$

其中 $\beta_N = \beta/N$, $x_{N+1} = x_1$.

正则系综平均的路径积分表示

- 上述表达式包含 **N 珠-环聚合物** (N -bead ring polymer) 的势能函数:

$$\mathcal{E}_N(\mathbf{x}) = \frac{1}{2\beta_N} \sum_{j=1}^N |x_j - x_{j+1}|^2 + \beta_N \sum_{j=1}^N V(x_j),$$

其中相邻的珠由弹性势能 $|x_j - x_{j+1}|^2$ 连接.

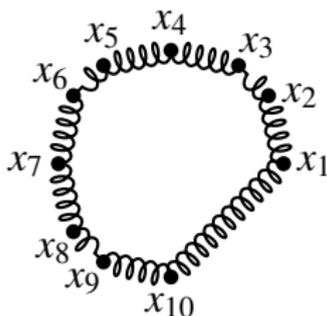


图: 一个 10 珠-环聚合物的图示



正则系综平均的路径积分表示

- 采样 \mathbb{R}^d 中的密度函数 $\langle x | e^{-\beta \hat{H}} | x \rangle$ 可转化为采样 \mathbb{R}^{dN} 中的经典 Gibbs 分布 $\pi_N(\mathbf{x}) \propto \exp(-\mathcal{E}_N(\mathbf{x}))$.

- 由于 Strang 分裂的局部误差是 $O(\beta_N^3)$, 使用

$$\langle \hat{O} \rangle_{\beta, N} := \frac{1}{\mathcal{Z}_N} \int_{\mathbb{R}^{dN}} \left(\frac{1}{N} \sum_{j=1}^N O(x_j) \right) \exp(-\mathcal{E}_N(\mathbf{x})) d\mathbf{x}$$

估计系综平均 $\langle \hat{O} \rangle_{\beta}$ 时的误差为 $O(N\beta_N^3)$.

- 路径积分表示只有在 N 足够大时才能得到精确的系综平均.



正则坐标下的 Langevin 动力学

利用正则坐标变换, 可以将 N 珠-环聚合物的势能函数改写为

$$\mathcal{E}_N(\boldsymbol{\xi}) = \frac{1}{2\beta_N^2} \sum_{k=0}^{N-1} \omega_k^2 |\xi_k|^2 + \beta_N \sum_{j=1}^N V \left(\underbrace{\sum_{k=0}^{N-1} \xi_k c_{j,k}}_{x_j(\boldsymbol{\xi})} \right),$$

其中 $c_{j,k}$ 为离散 Fourier 系数: $c_{j,0} = \frac{1}{\sqrt{\beta}}$,

$$c_{j,2k-1} = \sqrt{\frac{2}{\beta}} \sin\left(\frac{2\pi kj}{N}\right), \quad c_{j,2k} = \sqrt{\frac{2}{\beta}} \cos\left(\frac{2\pi kj}{N}\right), \quad k = 1, \dots, \frac{N-1}{2},$$

而 ω_k 对应于 N 珠-环聚合物的第 k 个模的特征振荡频率:

$$\omega_0 = 0, \quad \omega_{2k-1} = \omega_{2k} = \frac{2}{\beta_N} \sin\left(\frac{k\pi}{N}\right), \quad k = 0, 1, \dots, \frac{N-1}{2}.$$



正则坐标下的 Langevin 动力学

若在正则坐标 $\{\xi_k\}_{k=0}^{N-1}$ 的基础上再引入正则速度 $\{\eta_k\}_{k=0}^{N-1}$, 则也可得到欠阻尼版本的预条件 Langevin 动力学.

预条件欠阻尼 Langevin 动力学

$$\begin{cases} \dot{\xi}_k = \eta_k, \\ \dot{\eta}_k = -\xi_k - \frac{\beta_N}{\omega_k^2 + a} \sum_{j=1}^N \nabla V^a(x_j(\boldsymbol{\xi})) c_{j,k} - \eta_k + \sqrt{\frac{2}{\omega_k^2 + a}} \dot{B}_k, \end{cases} \quad k = 0, 1, \dots, N-1.$$

该动力学的不变分布是

$$\mu_N(\boldsymbol{\xi}, \boldsymbol{\eta}) \propto \exp \left(-\frac{1}{2} \sum_{k=0}^{N-1} (\omega_k^2 + a) (|\xi_k|^2 + |\eta_k|^2) - \beta_N \sum_{j=1}^N V^a \left(\sum_{k=0}^{N-1} \xi_k c_{j,k} \right) \right).$$



预条件 Langevin 动力学的遍历性

接下来叙述预条件 Langevin 动力学的遍历性结果. 需要如下假设:

基本假设

给定常数 $a > 0$, 势能函数

$$V^a(x) = V(x) - \frac{a}{2}|x|^2, \quad x \in \mathbb{R}^d,$$

在 \mathbb{R}^d 上二阶可微, 且存在常数 $M_1, M_2 \geq 0$ 使得:

- ① $V^a(x)$ 可以被分解为

$$V^a(x) = V^c(x) + V^b(x),$$

其中 $\nabla^2 V^c(x) \succ O_d$ 且 $|V^b(x)| \leq M_1$ 对任意 $x \in \mathbb{R}^d$ 成立.

- ② $V^a(x)$ 的 Hesse 矩阵满足

$$-M_2 I_d \preceq \nabla^2 V^a(x) \preceq M_2 I_d$$

对任意 $x \in \mathbb{R}^d$ 成立.



预条件 Langevin 动力学的遍历性

定理 3.2 (p. 41)

令 $(P_t)_{t \geq 0}$ 是预条件欠阻尼 Langevin 动力学的 Markov 半群, 则对 \mathbb{R}^{2dN} 中任意恒正的光滑函数 $f(\xi, \eta)$, 有

$$W_{\mu_N}(P_t f) \leq e^{-2\lambda_2 t} W_{\mu_N}(f), \quad \forall t \geq 0,$$

其中收敛速率 $\lambda_2 = \frac{a^2}{3M_2^2 + 5a^2} \exp(-4\beta M_1)$.

这里 $W_{\mu_N}(f)$ 是指密度函数 $f(\xi, \eta)$ 在 π_N 中的类相对熵:

$$W_{\mu_N}(f) = \left(\frac{M_2^2}{a^2} + 1 \right) \text{Ent}_{\mu_N}(f) + \sum_{k=0}^{N-1} \frac{1}{\omega_k^2 + a} \int_{\mathbb{R}^{2dN}} \frac{|\nabla_{\eta_k} f - \nabla_{\xi_k} f|^2 + |\nabla_{\eta_k} f|^2}{f} d\mu_N.$$



数值算例: 1 维势能函数

令势能函数 $V(x)$ 和观测函数 $O(x)$ 定义为:

$$V(x) = \frac{1}{2}x^2 + x \cos x, \quad O(x) = \sin\left(\frac{\pi}{2}x\right), \quad x \in \mathbb{R}.$$

采用预条件欠阻尼 Langevin 动力学, 取步长 $h = \frac{1}{16}$ 和演化时间 $T = 5 \times 10^6$.

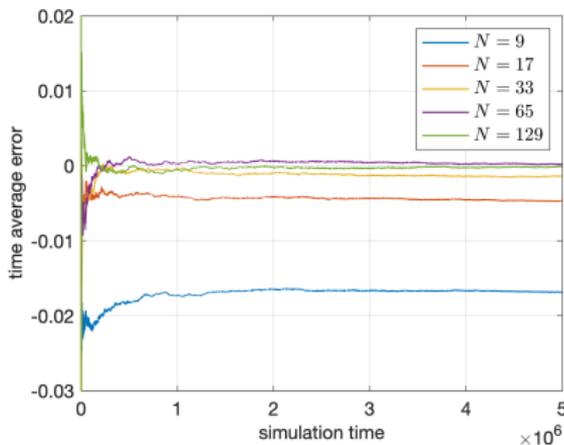
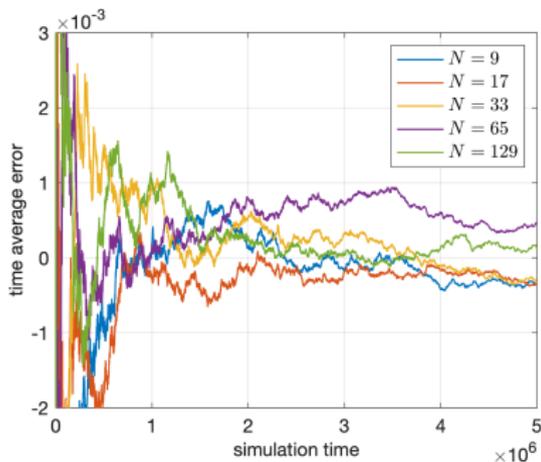


图: 计算正则系宗平均 $\langle O(x) \rangle_\beta$ 的时间平均误差. 左: $\beta = 1$. 右: $\beta = 4$.

可以看出不同的珠数 N 对应的收敛速度大致相同.



Section 3

随机分组方法



遍历性和采样误差

接下来叙述 RB-IPS 的遍历性和 discrete RB-IPS 的采样误差结果。

基本假设 1 (全局 Lipschitz 条件)

对于漂移力 $b(\cdot)$, 存在常数 L_0 使得

$$|b(x)| \leq L_0(|x| + 1), \quad \forall x \in \mathbb{R}^d.$$

对于相互作用力 $K(\cdot)$, 存在常数 L_1 使得

$$\max\{|K(x)|, |\nabla K(x)|, |\nabla^2 K(x)|\} \leq L_1, \quad \forall x \in \mathbb{R}^d.$$

全局 Lipschitz 条件保证了 Euler-Maruyama 离散格式的稳定性。



遍历性和长时间误差

基本假设 2 (漂移力条件)

给定 $r \geq 0$, 函数 $\kappa(r)$ 定义为

$$\kappa(r) = \inf \left\{ - \frac{(x-y) \cdot (b(x) - b(y))}{|x-y|^2} : x, y \in \mathbb{R}^d, |x-y| = r \right\},$$

且满足下列条件:

- ① $\kappa(r)$ 在 $(0, +\infty)$ 上连续;
- ② $\kappa(r)$ 在 $(0, +\infty)$ 上有一致的下界;
- ③ $\lim_{r \rightarrow \infty} \kappa(r) > 0$.

充分条件: 在 $\mathbb{R}^d \setminus B(0, R)$ 有 $\nabla^2 U(x) \succeq mI_d$ 对某个 $m > 0$ 成立.



遍历性和长时间误差

定理 4.5 (p. 78)

令 $(\tilde{Q}_n^h)_{n \geq 0}$ 为 discrete RB-IPS 的对偶半群. 存在常数 $L_1^*, h^* > 0$, 使得当

$$0 < L_1 \leq L_1^*, \quad 0 < h \leq h^*$$

时, 对于给定的初始分布 $\nu \in \mathcal{P}(\mathbb{R}^{dN})$, 有不等式:

$$\mathcal{W}_1(\nu \tilde{Q}_n^h, \pi) \leq C e^{-\lambda n h} + C \sqrt{h}.$$

常数 C, λ, L_1^*, h^* 仅依赖于 $d, \kappa(\cdot), \sigma, L_0$ 和初始分布 ν .

特别的, 这些常数都不依赖于粒子数 N , 分组大小 p 和步长 h .



遍历性和长时间误差

主要证明技术:

- 反射耦合方法

Eberle. (2016). Reflection couplings and contraction rates for diffusions. *Probability theory and related fields*, 166, 851-886.

- 三角不等式 (有限时间误差 + 遍历性 \implies 长时间误差)

Mattingly, Stuart & Tretyakov. (2010). Convergence of numerical time-averaging and stationary measures via Poisson equations. *SIAM Journal on Numerical Analysis*, 48(2), 552-577.

Schuh, & Souttar. (2024). Conditions for uniform in time convergence: applications to averaging, numerical discretisations and mean-field systems. arXiv preprint arXiv:2412.05239. (Inspired from my work)



数值算例: 相互作用系统

设外部势能 $U(x)$ 和相互作用势能 $V(x)$ 为

$$U(x) = \frac{1}{2}x^2 + 3.5e^{-(x-0.3)^2}, \quad V(x) = -e^{-x^2}.$$

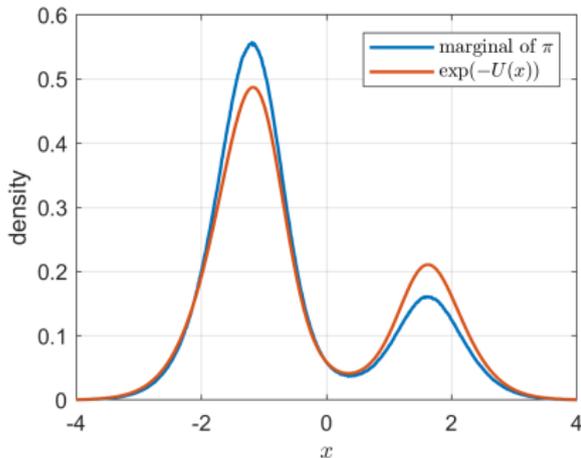


图: 目标分布 $\pi(\mathbf{x})$ 在 x^1 上的边缘分布与 $\exp(-U(x))$ 的对比, 反映了相互作用项 $V(x)$ 对整体分布的影响.



数值算例: 相互作用系统

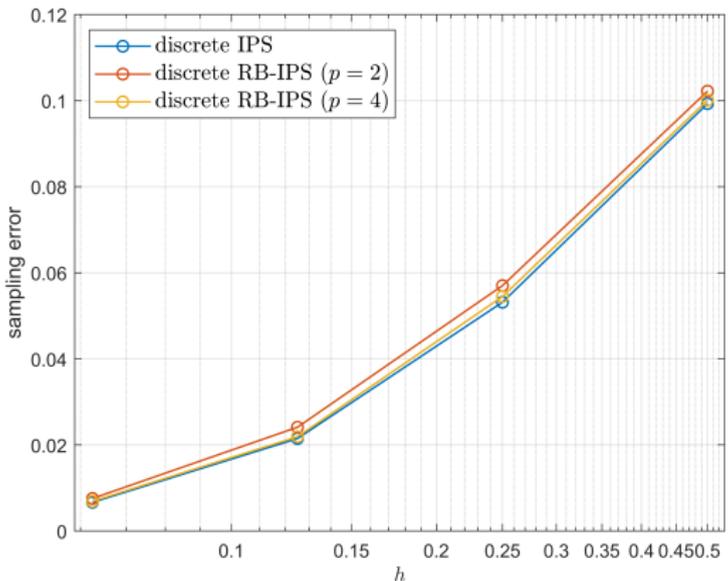


图: 当粒子数 $N = 64$ 时, discrete RB-IPS 在不同步长下的采样误差. 可以看出, 离散误差占主要部分, 而随机分组误差占比很小.

