Langevin Dynamics の多項式エルゴード性

Ergodic Lower Bounds

司馬博文

7/05/2024

Abstract

目標分布の裾が重ければ重いほど，Langevin 拡散過程の収束は遅くなる．本記事ではその様子を，平衡分布との全変動距離について，定量的に評価する．

$\mathbb{R}^n$ 上の Langevin 拡散を考える：

$$dX_t = -\nabla V(X_t)\,dt + \sqrt{2\beta^{-1}}\,dB_t, \qquad X_0 = x. \tag{1}$$

ただし，

$$V(x) = O(|x|^{2k}) \qquad (|x| \to \infty)$$

の仮定をおく．$k \geq 1/2$ の場合，指数エルゴード的であるが，$k < 1/2$ の場合はそうではない．

$k \in (0, 1/2)$ の設定で，次の ergodic lower bound を示したい：

$$C_1 \exp\left(c_1 V(x) - c_2 t^{\frac{k}{1-k}}\right) \leq \|P_t(x,-) - \mu_*\|_{\mathrm{TV}} \tag{2}$$

$$\mu_*(dx) \propto e^{-\beta V(x)} dx$$

この lower bound から，$k \in (0, 1/2)$ の場合，Langevin 過程 X が指数エルゴード的たり得ないことが従う．

式 (2) を示すためには，$G(x) := e^{\kappa V(x)}$ $(\kappa \in \mathbb{R})$ に対して，

$$\mathrm{E}_x[G(X_t)] \leq g(x,t)$$

を満たす関数 g を見つける必要がある (Hairer, 2021, pp. 34–35).

これは次の 3 ステップを辿る

1. そもそも $\mathrm{E}_x[G(X_t)] < \infty$ であることの証明（第 1 節）.
2. G に関するドリフト条件 $P_t\hat{L}G \leq C\varphi \circ G$ から，$\mathrm{E}_x[G(X_t)]$ の t に関する微分不等式を導く（第 2 節）.
3. 微分不等式から，Gronwall の補題より，結論を得る（第 3 節）.

1　$G = e^{\kappa V}$ の可積分性について

次元 $n = 1$ で考えてみる．
$V(x) = \frac{x^2}{2}$ とした場合，X は OU 過程になり，

$$\mathrm{E}_x[G(X_t)] < \infty \quad \Leftrightarrow \quad t < -\frac{1}{2}\log\left(1 - \frac{\beta}{\kappa}\right).$$

$V(x) = \log x$ とした場合，X は Bessel 過程になり，

$$\mathrm{E}_x[G(X_t)] < \infty \qquad (\forall_{t>0}).$$

$k \in (0, 1/2)$ の場合，∇V が有界であることに注目すれば，Bessel 過程の場合と同様に

$$\mathrm{E}_x[G(X_t)] < \infty \qquad (\forall_{t>0}).$$

1.1 はじめに

Markov 過程 X に関するドリフト条件

$$\hat{L}V \le -C\varphi \circ V \qquad \text{on } \mathbb{R}^n \setminus K$$

からは $V : E \to \mathbb{R}_+$ の可積分性が出る：

$$\mathrm{E}_x[V(X_t)] < \infty \qquad t \ge 0.$$

証明

上のドリフト条件を，(Hairer, 2021) の最も弱い意味で解釈すると

$$M_t := V(X_t) + C\int_0^t \varphi \circ V(X_s)\,ds$$

が任意の $x \in E$ に関して P_x-局所優マルチンゲールである，ということになる．
これだけの仮定でも，V が下に有界であるために M_t も下に有界であり，下に有界な局所優マルチンゲールは（真の）優マルチンゲールであることから，

$$\mathrm{E}_x\left[V(X_t) + C\int_0^t \varphi \circ V(X_s)\,ds\right] \le V(x).$$

加えて左辺が下に有界であることから，$\mathrm{E}_x[V(X_t)] < \infty$ でないと矛盾が起こる．

しかし，lower bound を得たい場合，

$$\hat{L}V \le C\varphi \circ G \qquad \text{on } \mathbb{R}^n \tag{3}$$

という情報のみから，

$$\mathrm{E}_x[G(X_t)] \le g(x,t) \; (< \infty)$$

という評価を得る必要が出てくる．この場合，$\mathrm{E}_x[G(X_t)] < \infty$ は非自明で，ケースバイケースの議論がである．

1.2 OU 過程の場合

An overdamped Langevin dynamics on $\mathbb{R}$ is defined as the solution to the following SDE:

$$dX_t = -\nabla V(X_t)\,dt + \sqrt{2\beta^{-1}}\,dB_t, \qquad X_0 = x_0.$$

If $V(x) = \frac{x^2}{2}$, X becomes an Ornstein-Uhlenbeck process. Transforming via $f(t,x) = xe^t$ and using Itô's formula, we get

$$X_t = x_0 e^{-t} + \sqrt{2\beta^{-1}}\int_0^t e^{-(t-s)}\,dB_s.$$

Hence, X is a Gaussian process with $X_t \sim \mathrm{N}\left(x_0 e^{-t}, \beta^{-1}(1-e^{-2t})\right)$.

In this case, expectation with respect to $G(y) = e^{\kappa V(y)} = e^{\frac{\kappa y^2}{2}}$ $(\kappa \in \mathbb{R})$ is given by

$$\begin{aligned}\mathrm{E}_x[G(X_t)] &= \int_{\mathbb{R}} G(y)\frac{1}{\sqrt{2\pi\beta^{-1}(1-e^{-2t})}}\exp\left(-\frac{(y-xe^{-t})^2}{2\beta^{-1}(1-e^{-2t})}\right)dy \\ &= \frac{1}{\sqrt{2\pi\beta^{-1}(1-e^{-2t})}}\int_{\mathbb{R}}\exp\left(\frac{\kappa\beta^{-1}(1-e^{-2t})y^2 - (y-xe^{-t})^2}{2\beta^{-1}(1-e^{-2t})}\right)dy.\end{aligned}$$

Taking a closer look at the numerator inside exp,

$$\begin{aligned}&\kappa\beta^{-1}(1-e^{-2t})y^2 - (y-xe^{-t})^2 \\ &= y^2\Big(\kappa\beta^{-1}(1-e^{-2t}) - 1\Big) - 2xe^{-t}y + x^2e^{-2t}.\end{aligned}$$

Therefore, we conclude

$$\mathrm{E}_x[G(X_t)] < \infty \quad \Leftrightarrow \quad \kappa\beta^{-1}(1-e^{-2t}) < 1.$$

In other words, $P_tG(x)$ is finite as long as

$$t < -\frac{1}{2}\log\left(1-\frac{\beta}{\kappa}\right).$$

1.3 Bessel 過程の場合

$V = a\log x$ ととると，$V'(x) = \frac{a}{x}$ であるから，これに関する Langevin 動力学は，$\beta = 1$ のとき，

$$dX_t = -\frac{a}{X_t}\,dt + dB_t$$

と，母数 a を持つ Bessel 過程になる．ただし，0 への到着時刻 T_0 で止めたもの X^{T_0} を考える．

> [@Lawler2019 p.10 命題 2.5]
>
> 母数 a を持つ Bessel 過程 X^{T_0} の密度を $q_t(x,y;a)$ で表す．このとき，
>
> $$q_t(x,y;1-a) = \left(\frac{y}{x}\right)^{1-2a} q_t(x,y;a)$$
>
> $$q_t(x,y;a) = q_t(y,x;a)\left(\frac{y}{x}\right)^{2a}$$
>
> $$q_{r^2t}(rx,ry;a) = \frac{1}{r}q_t(x,y;a)$$
>
> 加えて $a \geq \frac{1}{2}$ でもあるとき，
>
> $$q_1(x,y;a) = y^{2a}\exp\left(-\frac{x^2+y^2}{2}\right)h_a(xy),$$
>
> $$h_a(x) \sim \frac{1}{\sqrt{2\pi}}x^{-a}e^x \qquad (|x| \to \infty)$$

この結果は (Lawler, 2019, p. 59) をみる限り，修正 Bessel 関数と，Bessel 過程の Fokker-Planck 方程式との考察によって証明されている．

$$G(y) = e^{\kappa V(y)} = e^{a\kappa\log(y)} = y^{a\kappa}$$

であるから，密度 $q_t(x,y;a)$ に対してはどうやっても可積分である．

1.4 $k < 1/2$ の場合の尾部確率

$k < 1/2$ で最も大きく変わる点は，

$$\nabla V(x) = O(|x|^{2k-1}) \qquad (|x| \to \infty)$$

であるために，∇V が $\mathbb{R}^n$ 上で有界になることである．

このため，一般に SDE

$$dZ_t = b(Z_t)\,dt + \sigma(X_t)\,dB_t$$

の密度が，任意の $T > 0$ に対して，ある $A_T, a_T > a$ と $y \in \mathbb{R}$ が存在して

$$\frac{1}{A_T\sqrt{2\pi t}}e^{-\frac{a_T|y-x|^2}{2t}} \leq p_t(x,y) \leq \frac{A_T}{\sqrt{2\pi t}}e^{-\frac{|y-x|^2}{2a_Tt}}$$

$$t \in (0,T]$$

が成り立つことが使える．[1]

これによれば，

$$G(x) = e^{\kappa V(x)} = O(e^{\kappa|x|^{2k}}) \quad (|x| \to \infty)$$

に対して p_t の尾部が勝つため，$P_tG(x) < \infty$ である．

[1] (Kohatsu-Higa et al., 2022) で最初に知った．特に (Kohatsu-Higa, 2003) は詳しく扱っており，上からの評価は Malliavin 解析から得られる (Taniguchi, 1985). 同様にして熱方程式の基本解としても捉えられるが.

1.5 $k < 1/2$ の場合の G の可積分性

$k < 1/2$ の場合，式 (1) のドリフト係数が有界になる．このことから，G の可積分性が，X_t の密度の考察に依らず次のように導ける．

$$M := \max_{x\in\mathbb{R}^n} \nabla V(x)$$

と定める．$V(x) = O(|x|^{2k})$ ($|x| \to \infty$) より，ある $C > 0$ が存在して，

$$V(x) \le C|x|^{2k} \qquad \text{on } \mathbb{R}^n.$$

$$\begin{aligned} |X_t| &\le \int_0^t |\nabla V(X_t)|\,dt + \sqrt{2\beta^{-1}}|B_t| \\ &\le Mt + \sqrt{2\beta^{-1}}|B_t| \end{aligned}$$

より，

$$\begin{aligned} \mathrm{E}_x[|G(X_t)|] &\le \mathrm{E}_x\left[e^{\kappa V(|X_t|)}\right] \\ &\le \mathrm{E}_x\left[\exp\left(\kappa V(M_t + \sqrt{2\beta^{-1}|B_t|})\right)\right] \\ &\le e^{\kappa|Mt|^{2k}}\,\mathrm{E}_x\left[e^{\kappa 2^k\beta^{-k}|B_t|^{2k}}\right] < \infty. \end{aligned}$$

2 微分と拡張生成作用素の関係

(X_t) を $E = \mathbb{R}^n$ 上の Feller-Dynkin 過程，(P_t) をその遷移半群，$\widehat{L}$ をその拡張生成作用素とする．

命題 2

$G \in \mathcal{D}(\widehat{L})$ とする．すなわち，

$$t \mapsto M_t := G(X_t) - \int_0^t \widehat{L}G(X_s)ds$$

は任意の $x \in E$ について P_x-局所マルチンゲールである．
このとき，さらに G について次の条件を仮定する：

1. $\mathrm{E}_x[|G(X_t)|] < \infty$ $(x \in E, t \in \mathbb{R}_+)$．すなわち，$P_tG : E \to \mathbb{R}$ が定まる．
2. 同様に $\mathrm{E}_x[|\widehat{L}(G)(X_t)|] < \infty$ $(x \in E, t \in \mathbb{R}_+)$．すなわち，$\widehat{L}P_tG : E \to \mathbb{R}$ も定まる．[2]
3. $t \mapsto P_t\widehat{L}G(x)$ は局所有界．

このとき，$P_tG(x)$ は t で微分可能であり，次が導ける：

$$\frac{\partial}{\partial t}\mathrm{E}_x[G(X_t)] = \mathrm{E}_x[\widehat{L}G(X_t)].$$

これは，通常の意味での生成作用素 L の性質

$$\frac{\partial}{\partial t}P_tG = P_t(LG)$$

が，可積分性の条件の下で，拡張生成作用素 $\widehat{L}$ にも引き継がれると理解できる．

証明

仮定より，停止時の列 $\tau_n \nearrow \infty$ a.s. が存在し，任意の $n \in \mathbb{N}$ について，M^{τ_n} はマルチンゲールで，

$$\mathrm{E}_x\left[G(X_{t\wedge\tau_n}) - \int_0^{t\wedge\tau_n} \widehat{L}G(X_s)ds\right] = G(x), \qquad t \ge 0, x \in E. \tag{4}$$

仮定 1 より $\mathrm{E}_x[|G(X_{t\wedge\tau_n})|] < \infty$ であるから，

$$\mathrm{E}_x\left[\left|\int_0^{t\wedge\tau_n} \widehat{L}G(X_s)ds\right|\right] < \infty.$$

[2] 元々はある正の定数 $C > 0$ が存在して，$\widehat{L}G \le CG$．ある凹関数 φ について $\widehat{L}G \le \varphi\circ G$ が成り立つならばこの仮定は満たされることに注意，としていた．

でもある．従って Fubini-Tonelli の定理から

$$\mathrm{E}_x\left[\left|\int_0^{t\wedge\tau_n}\widehat{L}G(X_s)ds\right|\right]=\int_0^t\mathrm{E}_x\left[1_{[0,\tau_n]}(s)\widehat{L}G(X_s)\right]ds$$

と書き換えられる．
よって，式 (4) は

$$\mathrm{E}_x\left[G(X_{t\wedge\tau_n})\right]=G(x)+\int_0^t\mathrm{E}_x\left[1_{[0,\tau_n]}(s)\widehat{L}G(X_s)\right]ds$$

とも表せる．右辺が t について微分可能であるから，左辺も微分可能である：

$$\frac{\partial}{\partial t}\mathrm{E}_x\left[G(X_{t\wedge\tau_n})\right]=\mathrm{E}_x\left[1_{[0,\tau_n]}(t)\widehat{L}G(X_t)\right].$$

両辺の $n\to\infty$ に関する極限を取ると，右辺は $|\widehat{L}G(X_t)|$ が P_x-可積分であるから（仮定２），Lebesgue の優収束定理より，

$$\lim_{n\to\infty}\frac{\partial}{\partial t}\mathrm{E}_x[G(X_{t\wedge\tau_n})]=\lim_{n\to\infty}\mathrm{E}_x\left[1_{[0,\tau_n]}(t)\widehat{L}G(X_t)\right]=\mathrm{E}_x[\widehat{L}G(X_t)],\qquad x\in E,t\in(0,\infty).$$

加えてこの収束は，$t\in(0,\infty)$ に関して広義一様に起こる．実際，Hölder の不等式より，[3]

$$\begin{aligned}&\sup_{t\in[0,T]}\left|\frac{\partial}{\partial t}\mathrm{E}_x[G(X_{t\wedge\tau_n})]-\mathrm{E}_x[\widehat{L}G(X_t)]\right|\\&=\sup_{t\in[0,T]}\left|\mathrm{E}_x[1_{[0,\tau_n]}(t)\widehat{L}G(X_t)]-\mathrm{E}_x[\widehat{L}G(X_t)]\right|\\&=\sup_{t\in[0,T]}\left|\mathrm{E}_x\left[(1-1_{[0,\tau_n]}(t))\widehat{L}G(X_t)\right]\right|\\&\le\sup_{t\in[0,T]}\mathrm{E}_x\left[(1-1_{[0,\tau_n]}(T))|\widehat{L}G(X_t)|\right]\\&\le\|1-1_{[0,\tau_n]}(T)\|_{L^\infty(\Omega)}\sup_{t\in[0,T]}\mathrm{E}_x\left[|\widehat{L}G(X_t)|\right]\xrightarrow{n\to\infty}0.\end{aligned}$$

最後の不等式にて，仮定３による局所有界性

$$\sup_{t\in[0,T]}\mathrm{E}_x\left[|\widehat{L}G(X_t)|\right]<\infty$$

を用いた．
この導関数の一様収束と，Lebesgue の優収束定理による各点収束

$$\mathrm{E}_x[G(X_{t\wedge\tau_n})]\xrightarrow{n\to\infty}\mathrm{E}_x[G(X_t)]$$

を併せると，$\mathrm{E}_x[G(X_t)]$ も可微分で，その導関数は極限

$$\frac{\partial}{\partial t}\mathrm{E}_x[G(X_t)]=\lim_{n\to\infty}\frac{\partial}{\partial t}\mathrm{E}_x[G(X_{t\wedge\tau_n})]=\mathrm{E}_x[\widehat{L}G(X_t)]$$

として得られることが結論づけられる．

[@Rudin-Principles p.152 定理 7.17][a]

[a] [@杉浦光夫 1980 p.311] 定理 13.7 系では，f_n に C^1-級の仮定を置いて，この場合は f が C^1-級になることを導いている．

$f_n:[a,b]\to\mathbb{R}$ を可微分な関数列とし，ある関数 $f:[a,b]\to\mathbb{R}$ に各点収束するものとする．
仮に，導関数列 $\{f_n'\}$ が一様位相に関して Cauchy 列ならば，$f_n\to f$ も一様収束し，加えて f も可微分で，

$$\lim_{n\to\infty}f_n'(x)=f'(x)$$

[3] $\sup_{t\in[0,T]}\widehat{L}G(X_t)$ は可積分とは限らないため，sup を期待値の中に入れることはできない．Hölder の不等式により，これを迂回できる．

が成り立つ．

3 下界の導出

元来の目的である下界の導出のためには，

$$\mathrm{E}_x[G(X_t)] \le C G(x) \exp\left(c t^{\frac{k}{1-k}}\right)$$

という評価を得る必要がある．Gronwall の不等式を用いれば，導関数に関する不等式

$$\frac{\partial}{\partial t} \mathrm{E}_x[G(X_t)] \le \mathrm{E}_x[\hat{L} G(X_t)] \le C\, \mathrm{E}_x[\varphi \circ G(X_t)]$$

があれば十分である．この導関数に関する不等式は，命題 2 とドリフト条件 (3)

$$\hat{L} G \le C \varphi \circ G$$

を併せることで，

$$\frac{\partial}{\partial t} \mathrm{E}_x[G(X_t)] = \mathrm{E}_x[\hat{L} G(X_t)] \le C\, \mathrm{E}_x[\varphi \circ G(X_t)]$$

より得る．

参考文献

Hairer, M. (2021). *Convergence of markov processes*.

Kohatsu-Higa, A. (2003). Lower bounds for densities of uniformly elliptic non-homogeneous diffusions. In E. Giné, C. Houdré, and D. Nualart, editors, *Stochastic inequalities and applications*, pages 323–338. Basel: Birkhäuser Basel.

Kohatsu-Higa, A., Nualart, E., and Tran, N. K. (2022). Density estimates for jump diffusion processes. *Applied Mathematics and Computation*, *420*, 126814.

Lawler, G. F. (2019). *Notes on the bessel process*.

Taniguchi, S. (1985). Applications of Malliavin's calculus to time-dependent systems of heat equations. *Osaka Journal of Mathematics*, *22*(2), 307–320.