STAT0041: Stochastic Calculus

Lecture 14 - Ornstein–Uhlenbeck Process and Langevin Dynamics

Lecturer: Weichen Zhao Fall 2024

Key concepts:

- *Ornstein–Uhlenbeck semigroup;*
- *Langevin dynamics.*

14.1 Ornstein–Uhlenbeck Process

Recall Ornstein–Uhlenbeck processes

$$dX_t = -bX_t dt + \sigma dB_t X_0,$$

which solution is

$$X_t = X_0 e^{-bt} + \sigma e^{-bt} \int_0^t e^{bs} dB_s$$

In this lecture we consider following specific OU process

$$dX_t = -X_t dt + \sqrt{2} dB_t. \tag{14.1}$$

which solution is

$$\begin{aligned} X_t &= e^{-t} X_0 + \sqrt{2} e^{-t} \int_0^t e^s dB_s \\ &= e^{-t} X_0 + e^{-t} B_{e^{2t}-1} \\ &= e^{-t} X_0 + \sqrt{1 - e^{-2t}} z \qquad z \sim N(0,1) \\ &\sim N(e^{-t} X_0, 1 - e^{-2t}) \end{aligned}$$

Ornstein–Uhlenbeck Operator. The generator of OU process (14.1) is

$$\mathcal{L}_{OU} f = -x \cdot \nabla f + \Delta f,$$

and adjoint generator is

$$\mathcal{L}_{OU}^* g = \nabla \cdot (xg) + \Delta g$$

Solving equation $\mathcal{L}_{OU}^* \pi = 0$ we have the invariant measure

$$d\pi(x) \propto e^{-\frac{1}{2}x^2} dx$$

of OU process (14.1), which is the standard Gaussian measure on $\mathbb{R}^n$.

Ornstein–Uhlenbeck Semigroup. From the solution of OU process (14.1)

$$X_t = e^{-t} X_0 + \sqrt{1 - e^{-2t}} z \qquad z \sim N(0, 1)$$

we have the transition function

$$p(t, x, y) = \frac{1}{\sqrt{2\pi(1 - e^{-2t})}} \exp\left(-\frac{(y - xe^{-t})^2}{2(1 - e^{-2t})} \right).$$

Then the transition semigroup is

$$\begin{aligned} P_t f(x) &= \mathbb{E}[f(X_t) \mid X_0 = x] = \int f(y) p(t, x, dy) \\ &= \mathbb{E}[f(e^{-t} x + \sqrt{1 - e^{-2t}} z)] \qquad z \sim N(0, 1) \\ &= \int f(e^{-t} x + \sqrt{1 - e^{-2t}} z) d\pi(z) \end{aligned}$$

which is called *Ornstein–Uhlenbeck Semigroup*

Kolmogorov Equations

Kolmogorov backward equation:

$$\frac{\partial}{\partial t} P_t f(x) = \mathcal{L}_{OU} P_t f(x) = -x \cdot \nabla P_t f(x) + \Delta P_t f(x)$$

Fokker-Planck equation:

$$\partial_t \mu(x, t) = \mathcal{L}_{OU}^* \mu(x, t) = \nabla \cdot (x\mu(x, t)) + \Delta \mu(x, t)$$

14.2 Langevin Dynamics

Definition 14.1 (Langevin Dynamics) *Given potential function $V(x)$, Langevin dynamics is the following SDE*

$$\mathrm{d}X_t = -\nabla V(X_t)\mathrm{d}t + \sqrt{2}\mathrm{d}B_t. \tag{14.2}$$

which solution is called Langevin Diffusion.

The generator of Langevin diffusion is

$$\mathcal{L}_{LD} f = -\nabla V \cdot \nabla f + \Delta f,$$

and adjoint generator is

$$\mathcal{L}_{LD}^* g = \nabla \cdot (g \nabla V) + \Delta g$$

Kolmogorov backward equation:

$$\frac{\partial}{\partial t} P_t f(x) = \mathcal{L}_{LD} P_t f(x) = -\nabla V(x) \cdot \nabla P_t f(x) + \Delta P_t f(x)$$

Fokker-Planck equation:

$$\partial_t \mu(x,t) = \mathcal{L}_{LD}^* \mu(x,t) = \nabla \cdot (\mu(x,t) \nabla V(x)) + \Delta \mu(x,t)$$

Proposition 14.2 *The invariant measure of Langevin diffuison*

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

is

$$d\pi(x) \propto e^{-V(x)} dx.$$

Remark 14.3 *In fact, for all given $p(x)$, we can write it as*

$$p(x) = e^{-\log(-p(x))}.$$

It suggests that we can construct Langevin diffusion with invariant measure $p(x)$ by well potential function $V(x)$. That is we can generate sample $X_T \sim p(x)$ from sample X_0 which law is an arbitrary distribution by iterating through Langevin dynamics.

Besides being concerned with the form of the invariant measure of Langevin diffusion, another issue of great interest to us is the rate at which it converges to invariant measure. We begin with giving the mathematical quantity that characters the distance between distributions

Definition 14.4 (Wasserstein distance) *2-Wasserstein distance between probability measures μ and ν is defined as*

$$W_2(\mu, \nu) := \inf_{\gamma \in \mathcal{C}(\mu,\nu)} \left(\int \|x - y\|^2 \gamma(\mathrm{d}x, \mathrm{d}y) \right)^{\frac{1}{2}}. \tag{14.3}$$

where $\mathcal{C}(\mu, \nu)$ is the set of all couplings of μ and ν. We say γ is a coupling of μ and ν, if its marginal on the first variable is μ and its marginal on the second is ν.

We say a measure μ is α-strongly log-concave if $\mu \propto e^{-V}$ with V being α-strongly convex, that is

$$\nabla^2 V \succeq \alpha I.$$

Theorem 14.5 *Let $\{X_t\}$ be generated according to the Langevin diffusion with initialization $X_0 \sim \mu_0$ and stationary measure $\mu \propto e^{-V}$. Assume μ is α-strongly logconcave, then*

$$W_2^2(\mu_t, \mu) \leq \exp(-2\alpha t) W_2^2(\mu_0, \mu).$$