STAT0041: Stochastic Calculus

Lecture 14 - Ornstein–Uhlenbeck Process and Langevin Dynamics

Lecturer: Weichen Zhao　　Fall 2025

Key concepts:

- *OU*过程；
- *Langevin*扩散。

14.1 Ornstein–Uhlenbeck过程

设$(\Omega, \mathscr{F}, (\mathscr{F}_t), \mathrm{P})$为一个带流的概率空间，$B_t$ 为一个$\mathscr{F}_t$-布朗运动，回顾OU过程

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

我们计算过它的解为

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

本节我们考察一个具体地OU过程，

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

那么它的解为

$$\begin{aligned}
X_t &= e^{-t} X_0 + \sqrt{2} e^{-t} \int_0^t e^s dB_s \\
&\overset{d}{=} e^{-t} X_0 + e^{-t} B_{e^{2t}-1} && \text{(等号在分布意义下成立)} \\
&\overset{d}{=} e^{-t} X_0 + \sqrt{1 - e^{-2t}} z && \text{(令}z \sim N(0,1)\text{)} \\
&\sim N(e^{-t} X_0, 1 - e^{-2t})
\end{aligned}$$

验证第二个等号：设 $Y_t \triangleq \sqrt{2} \int_0^t e^s \, dB_s$，由于$e^s$不具有随机性，由Itô积分的定义以及布朗运动增量的高斯性，Y_t服从高斯分布。又Itô积分的均值为0，我们只需要计算二阶矩。

由Itô等距

$$\mathbb{E}[Y_t^2] = 2\int_0^t (e^s)^2\, ds = 2 \cdot \frac{1}{2}\, e^{2s}\big|_0^t = e^{2t} - 1$$

因此，随机变量 Y_t 的分布为：

$$Y_t \sim \mathcal{N}\left(0, e^{2t} - 1\right) \stackrel{d}{=} B_{e^{2t}-1}.$$

OU算子与不变测度. 我们已经证明Itô扩散过程

$$dX_t = b(X_t)dt + \sigma(X_t)dB_t$$

的生成元为

$$\begin{aligned}\mathcal{L}f &= \sum_i b_i(x)\frac{\partial}{\partial x_i}f + \frac{1}{2}\sum_{i,j}\Sigma_{ij}(x)\frac{\partial^2}{\partial x_i \partial x_j}f \\ &= b(x)\cdot\nabla f + \frac{1}{2}Tr(\Sigma^T(x)\nabla^2 f)\end{aligned}$$

对于OU过程$dX_t = -X_t dt + \sqrt{2}dB_t$，其生成元OU算子为

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

其伴随算子为

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

解方程$\mathcal{L}_{OU}^*\pi = 0$，可以得到OU过程的不变测度

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

这即是$\mathbb{R}^n$上的标准高斯测度$N(0,1)$。

转移函数与OU半群. 从OU过程的解

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

可以得到OU过程的转移函数

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

那么转移半群为

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

称为OU半群。

Kolmogorov方程. OU过程的Kolmogorov向后方程：

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

Fokker-Planck方程：

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

14.2 Langevin Dynamics

Definition 14.1 (Langevin Dynamics) *给定能量函数 $V(x)$，Langevin Dynamics 是如下形式的SDE*

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

其解一般称为 Langevin 扩散。

Langevin 扩散的生成元为

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

生成元的伴随为

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

Kolmogorov向后方程：

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

Fokker-Planck方程：

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

Proposition 14.2 *Langevin* 扩散

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

的不变测度为

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

Proof: Langevin 扩散的不变测度π满足方程：

$$\begin{aligned}
&\nabla \cdot (\pi \nabla V(x)) + \Delta\pi = 0 \\
\iff\ &\nabla \cdot (\pi \nabla V(x)) + \nabla \cdot \nabla\pi = 0 \\
\iff\ &\nabla \cdot (\pi \nabla V(x) + \nabla\pi) = 0 \\
\iff\ &\nabla \cdot (\pi \nabla V(x) + \pi\nabla \log(\pi)) = 0 \\
\iff\ &V(x) + \log(\pi) = 0
\end{aligned}$$

即

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

■

注. 事实上，对于任意给定连续分布$p(x)$，我们可以把它写成

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

这表明通过适当的选取能量函数$V(x)$，我们可以构造 Langevin 扩散使得它的不变测度是目标分布$p(x)$，也就是说，通过Langevin Dynamics的迭代，我们可以从服从任意分布的样本X_0产生服从目标分布$p(x)$的样本X_T。

我们除了关心Langevin 扩散转移到最终的平稳分布的形式，另一个非常关心的问题是它收敛到平稳分布的速度，首先给出刻画分布之间距离的数学量。

Definition 14.3 (Wasserstein距离) 概率测度 μ 和 ν 之间的*2-Wasserstein* 距离定义为：

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

其中$\mathcal{C}(\mu,\nu)$是μ和ν的所有耦合*(Couplings)*构成的空间，$\|\cdot\|$是欧式范数。

注1. 我们称$\gamma(x,y)$是测度$\mu(x)$和$\nu(y)$的耦合，如果它关于x的边缘分布是μ，关于y的边缘分布是ν。

注2. 我们也可以把W_2推广到W_p

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

Condition 14.4 设$\alpha > 0$为一个常数，称一个测度

$$\mu \propto e^{-V}$$

是α-强*log-concave*的，如果V是α-强凸的，即

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

Theorem 14.5 设 X_t 为初值为 $X_0 \sim \mu_0$，平稳分布为 $\mu \propto e^{-V}$ 的*Langevin*扩散，假设μ是α-强*log-concave*的，那么

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

Proof: 设γ_0是(μ_0,μ)的最优耦合(最优耦合是存在的)，即

$$\left(\int \|x-y\|^2 \gamma_0(\mathrm{d}x,\mathrm{d}y)\right)^{\frac{1}{2}} = W_2(\mu_0,\mu)$$

设$(X_0,X_0^*) \sim \gamma_0$，即$X_0 \sim \mu_0$，$X_0^* \sim \mu$，随Langevin Dynamics演化$t$时间，有以下两个Dynamics：

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

记$(X_t,X_t^*) \sim \gamma_t$，$X_t \sim \mu_t$，$X_t^* \sim \mu$，则$\gamma_t$是$(\mu_t,\mu)$的耦合。我们的目标是要控制$\mathbb{E}_{(X_t,X_t^*)\sim\gamma_t}[\|X_t - X_t^*\|^2]$。

$$\begin{aligned}
&d\|X_t - X_t^*\|^2 \\
&= 2\langle X_t - X_t^*, dX_t - dX_t^*\rangle \\
&= 2\langle X_t - X_t^*, -\nabla V(X_t)dt + \nabla V(X_t^*)dt\rangle \qquad \text{(同一个B.M.驱动)} \\
&= -2\langle X_t - X_t^*, \nabla V(X_t) - \nabla V(X_t^*)\rangle dt \\
&= -2\langle X_t - X_t^*, \nabla^2 V((1-\tau)X_t + \tau X_t^*)(X_t - X_t^*)\rangle dt \qquad \text{(中值定理)} \\
&\le -2\alpha \|X_t - X_t^*\|^2 dt
\end{aligned}$$

两边积分有

$$\|X_t - X_t^*\|^2 \leq \exp(-2\alpha t)\|X_0 - X_0^*\|^2$$

取期望

$$\begin{aligned} W_2^2(\mu_t, \mu) &\leq \mathbb{E}_{\gamma_t}[\|X_t - X_t^*\|^2] && (W_2\text{是最优耦合}) \\ &\leq \exp(-2\alpha t)\mathbb{E}_{\gamma_0}\|X_0 - X_0^*\|^2 \\ &= \exp(-2\alpha t)W_2^2(\mu_0, \mu) && (\gamma_0\text{是最优耦合}) \end{aligned}$$

■