[AI2613 Lecture 21] Markov Semigoup

May 28, 2024

1 Review of Itô Lemma

Recall that in the previous lectures, we formalized the diffusion $\{X_t\}_{t\geq 0}$ as a stochastic differential equation

$$\mathrm{d}X_t = \mu(t, X_t)\,\mathrm{d}t + \sigma(t, X_t)\,\mathrm{d}B_t$$

where $\{B_t\}_{t\geq 0}$ is a standard Brownian motion. The meaning of the differential equation is interpreted as:

$$\forall T \geq 0,\ X_T = X_0 + \int_0^T \mu(t, X_t)\,\mathrm{d}t + \int_0^T \sigma(t, X_t)\,\mathrm{d}B_t$$

where $\int_0^T \mu(t, X_t)\,\mathrm{d}t$ is the ordinary Riemann integral of $\mu(t, X_t)$ and $\int_0^T \sigma(t, X_t)\,\mathrm{d}B_t$ is the Itô Integral of $\sigma(t, X_t)$. The Itô integral $\int_0^T X_t\,\mathrm{d}B_t$ is defined as the mean square limit of

$$\sum_{i=1}^n X_{t_{i-1}}(B_{t_i} - B_{t_{i-1}})$$

as $n \to \infty$. In the last lecture, we proved that

$$\lim_{n\to\infty} \sum_{i=1}^n (B_{t_i} - B_{t_{i-1}})^2 = \int_0^T (\mathrm{d}B_t)^2 = T = \int_0^T \mathrm{d}t$$

holds for any T. This indicates that $(\mathrm{d}B_t)^2 \approx \mathrm{d}t$ holds in the limiting case(the mean square limit). In most cases, we can replace $(\mathrm{d}B_t)^2$ with $\mathrm{d}t$ to simplify the calculation. With this property, we showed the chain rule of the Itô integral (Itô Lemma)

$$\begin{aligned}
\mathrm{d}f(X_t) &= f'(X_t)\,\mathrm{d}X_t + \frac{1}{2}f''(X_t)(\mathrm{d}X_t)^2 + o(\mathrm{d}t) \\
&= \left(f'(X_t)\mu(t, X_t) + \frac{1}{2}f''(X_t)(\sigma(t, X_t))^2\right)\mathrm{d}t + f'(X_t)\sigma(t, X_t)\,\mathrm{d}B_t
\end{aligned}$$

2 Markov Semigroup

For a stochastic process, there are two ways to describe it: one is to describe the local walk for the random variable at each step (e.g., using the stochastic differential equation) and the other is to introduce the transition matrix.In this section, we will introduce the "transition operators" for the diffusion.

Recall that the transition matrix $P = (P_{ij})_{i,j\in[n]}$ of the finite Markov chain is defined as follwos:

$$P(i, j) = \mathbf{Pr}\left[X_{t+1} = j \mid X_t = i\right].$$

And we have that

$$P^t(i,j) = \mathbf{Pr}\left[X_{s+t} = j | X_s = i\right],$$

where $P^t(i,j)$ represents the probability of X_s walks from i to j in exactly t steps. By the law of total probability (or the definition of matrix multiplication), we have that

$$P^{s+t}(i,j) = (P^s \cdot P^t)(i,j) = \sum_{k\in[n]} P^s(i,k)P^t(k,j) \tag{1}$$

If π_t is the probability distribution of X_t, then we have that $\pi_t^\top P = \pi_{t+1}^\top$.

Another way to characterize P is to view it as an *operator*. Let $f\colon [n] \to \mathbb{R}$ be an arbitrary function. We can equivalently view it as an vector in $f = (f(1), \ldots, f(n)) \in \mathbb{R}^n$. If we multiply f to the right of P, by definition, we have that

$$[Pf](i) = \sum_{j\in[n]} P(i,j)f(j) = \mathbf{E}\left[f(X_{t+1}) \mid X_t = i\right]. \tag{2}$$

As a result, Pf is another function on $[n]$ whose value at the input i is $\mathbf{E}\left[f(X_{t+1}) \mid X_t = i\right]$. The value has the intuitive interpretation that $[Pf](i)$ is the average of the values $f(j)$, where j is the next position in the random walk P when one is currently at i.

In the same way, we have that

$$[P^t f](i) = \mathbf{E}\left[f(X_{s+t}) \mid X_s = i\right]. \tag{3}$$

We use this way to describe the "transition matrix" of a diffusion. Note that now the function f is not necessarily of finite dimension.

Following the same property 1-3, we can define the "transition operators" that is similar to $\{P^t\}_{t\ge 0}$ for the diffusion model.

Definition 1 (Markov Semigroup)**.** *Let $(E, \mathcal{F})$ be a measurable space. Let $\boldsymbol{Q} = \{Q_t\}_{t\ge 0}$ be a family of Markov operators on $(E,\mathcal{F})$ which satisfies $Q_t : E \mapsto (\mathcal{F} \mapsto \mathbb{R})$. Then $\boldsymbol{Q}$ is a* ***Markov semigroup*** *on $(E,\mathcal{F})$ if and only if*

1. *$\forall t \ge 0, \forall x \in E$, $Q_t(x)$ is a measure on $(E,\mathcal{F})$*
2. *$\forall x \in E$, $Q_0(x,\ \mathrm{d}y) = \delta_x(\mathrm{d}y)$*
3. *$\forall s,t \ge 0, \forall x \in E$, $Q_{s+t}(x, A) = \int_E Q_s(x,\ \mathrm{d}y)Q_t(y,A)$*

Here δ_x is the Dirac delta function which satisfies:

$$\delta_x(t) \simeq \begin{cases} \infty & ,t = x \\ 0 & ,t \ne x \end{cases}, \int_{-\infty}^{+\infty} \delta_x(t)\, dt = 1$$

Note that condition 3 in the definition of Markov semigroup corresponds to equation 1 about the transition matrix of the finite Markov chain. Also, Q_t can be considered as an operator that maps functions to functions. We define $Q_t(f)$ as follows:

$$Q_t f(x) = \int_{y\in E} f(y) Q_t(x,\ \mathrm{d}y)$$

which is the integration version of equation 2. Now we can define the Markov chain using the Markov Semigroup.

Definition 2 (Markov chain with Markov Semigroup)**.** *We say a continuous-time Markov chain $\{X_t\}_{t\geq 0}$ is attained by the Markov semigroup Q on $(E, \mathcal{F})$ if and only if the following holds:*

$$\forall t, \forall f \in C_0(E), Q_t f(x) = \mathbf{E}\left[f(X_t) \mid X_0 = x\right]$$

Here, $C_0(E)$ is the set of continuous functions tending to zero at infinity (in this class, we only consider the functions in $C_0(E)$). Also, we require that

1. $\forall t \geq 0, \forall f \in C_0(E)$, $Q_t(f) \in C_0(E)$.
2. $\forall f \in C_0(E)$, $\|Q_t f - f\|_\infty \to 0$ as $t \to 0$.

A semigroup Q satisfying these additional conditions is a Feller semigroup. We will only study the Feller semigroup in the lecture unless otherwise specified.

By definition of the Markov Semigroup, we can derive the following proposition, which also corresponds to equation 1 .

Proposition 3 (Chapman-Kolmogorov equation)**.** *Q is a Markov semigroup on $(E, \mathcal{F})$, then for any $f \in C_0(E)$, we have that*

$$Q_{s+t}(f) = Q_s Q_t(f)$$

That is, $Q_{s+t} = Q_s Q_t$.

Proof. By definition, we have that

$$\begin{aligned}
Q_{s+t} f(x) &= \int_{y \in E} f(y) Q_{s+t}(x, \mathrm{d}y) \\
&= \int_{y \in E} f(y) \int_{z \in E} Q_t(z, \mathrm{d}y) Q_s(x, \mathrm{d}z) \\
&= \int_{z \in E} (\int_{y \in E} f(y) Q_t(z, \mathrm{d}y)) Q_s(x, \mathrm{d}z) \\
&= \int_{z \in E} (Q_t f(z)) Q_s(x, \mathrm{d}z) = Q_s(Q_t f(x))
\end{aligned}$$

The third equation follows from the Fubini's theorem. □

3 *Generator of the Markov Semigoup*

For a continuous-time Markov chain, we cannot expect a notion identical to the transition matrix for a discrete-time Markov chain. Nevertheless, the following notion of *generator*, is in a sense similar to $P - I$ for the transition matrix P and identity matrix I.

Definition 4 (Generator of the Markov Semigoup)**.** *Assume that Q is a Markov semigroup on $(E, \mathcal{F})$. The operator $\mathcal{L}$ is the* ***generator*** *of Q if and only if*

$$\mathcal{L} f = \lim_{t \downarrow 0} \frac{Q_t f - t}{t}$$

for those $f \in D(\mathcal{L}) := \{f \in C_0(E) \ : \ \text{the above limit exists}\}$.

We have that $\mathcal{L} = \frac{\mathrm{d}}{\mathrm{d}t} Q_t\big|_{t=0}$ and therefore $\mathcal{L}$ is the "rate" of change under the family of operators $\{Q_t\}_{t\geq 0}$.

Note that the Markov semigroup corresponds to the set of the power of the transition matrix P: $\{P^t\}_{t\geq 0}$. As with the finite Markov chain, Q_t also can be described by $\mathcal{L}$.

Proposition 5. *$Q_t = e^{t\mathcal{L}}$ holds for any $t \geq 0$.*

Note that the exponential of the operator A is

$$e^A := \sum_{k\geq 0} \frac{A^k}{k!}$$

Proof. By Proposition 3, for any $f \in C_0(E)$, we have that

$$\begin{aligned}\mathcal{L}(Q_t f) &= \lim_{s\downarrow 0} \frac{Q_s(Q_t f) - Q_t f}{s} = \lim_{s\downarrow 0} \frac{Q_{t+s} f - Q_t f}{s} \\ &= \left(\lim_{s\downarrow 0} \frac{Q_{t+s} - Q_t}{s}\right) f = \left(\frac{\mathrm{d}}{\mathrm{d}t} Q_t\right) f\end{aligned}$$

The solution of the differential equation $\mathcal{L}Q_t = \frac{\mathrm{d}}{\mathrm{d}t} Q_t$ is $Q_t = e^{t\mathcal{L}}$. □

One useful property of $\mathcal{L}$ is that we can commute $\mathcal{L}$ and Q_t.

Proposition 6. *$\mathcal{L}$ and Q_t are commutative for any $t \geq 0$.*

Proof. By definition, we have that

$$\begin{aligned}\mathcal{L}(Q_t f) &= \lim_{s\downarrow 0} \frac{Q_s Q_t f - Q_t f}{s} = \lim_{s\downarrow 0} \frac{Q_t Q_s f - Q_t f}{s} \\ &= \lim_{s\downarrow 0} Q_t \frac{Q_s f - f}{s} = Q_t \lim_{s\downarrow 0} \frac{Q_s f - f}{s} = Q_t \mathcal{L} f\end{aligned}$$

The second equation holds since by Proposition 3, we have that

$$Q_s Q_t = Q_{s+t} = Q_{t+s} = Q_t Q_s$$

The third and fourth equations hold since Q_s is a continuous linear operator for any $s \geq 0$. □

Then we see an example of the generator.

Example 1 (The generator of the Poisson process)**.** *In the Poisson process, $X_t - X_0 \sim \mathrm{Poisson}(\lambda t)$, that is*

$$\mathbf{Pr}\left[X_t - X_0 = k\right] = \frac{(\lambda t)^k}{k!} e^{-\lambda t}$$

By definition of generator and Markov semigroup, we have that

$$\begin{aligned}\mathcal{L} f(n) &= \lim_{t\downarrow 0} \frac{\mathbf{E}\left[f(X_t) \mid X_0 = n\right] - f(n)}{t} \\ &= \lim_{t\downarrow 0} \frac{1}{t} \left(\sum_{k=0}^{\infty} \frac{(\lambda t)^k}{k!} e^{-\lambda t} f(n+k)\right) - \frac{f(n)}{t} \\ &= \lim_{t\downarrow 0} \frac{e^{-\lambda t}}{t} (f(n) + \lambda t f(n+1) + o(t)) - \frac{f(n)}{t} \\ &= \lim_{t\downarrow 0} \lambda e^{-\lambda t} f(n+1) + \frac{(1 - \lambda t + o(t) - 1)}{t} f(n) \\ &= \lambda(f(n+1) - f(n))\end{aligned}$$

The fourth equation follows from the taylor expansion of $e^{-\lambda t}$.

The above example implies that the generator of the Poisson process with rate λ is a infinite-dimensional matrix with diagonal elements of $-\lambda$ and superdiagonal elements of λ. Based on Proposition 5, we can write out the Markov semigroup of the Poisson process. This indicates that we can use the above generator to give another definition for the Poisson process.

For the Itô process, we can also derive the generator of it by Itô Lemma.

Lemma 7. *For the diffusion $\{X_t\}_{t\geq 0}$ on $(E, \mathcal{F})$ which satisfies that*

$$\mathrm{d}X_t = \mu(X_t)\,\mathrm{d}t + \sigma(X_t)\,\mathrm{d}B_t$$

The generator of $\{X_t\}_{t\geq 0}$ is as follows

$$\forall f \in C_0(E), \mathcal{L}f = f'\mu + \frac{1}{2}f''\sigma^2$$

Proof. By definition of the generator, we have that

$$\begin{aligned}
\mathcal{L}f(x) &= \lim_{t\downarrow 0} \frac{\mathbf{E}\left[f(X_t) \mid X_0 = x\right] - f(x)}{t} \\
&= \lim_{t\downarrow 0} \frac{\mathbf{E}\left[f(X_t) - f(X_0) \mid X_0 = x\right]}{t} \\
&= \mathbf{E}\left[\frac{\mathrm{d}}{\mathrm{d}t} f(X_t)|_{t=0} \,\middle|\, X_0 = x\right]
\end{aligned}$$

We apply the Itô Lemma for $\mathrm{d}f(X_t)$,

$$\begin{aligned}
\mathcal{L}f(x) &= \frac{1}{\mathrm{d}t}\mathbf{E}\left[(f'(x)\mu(x) + \frac{1}{2}f''(x)(\sigma(x))^2)\,\mathrm{d}t + f'(x)\sigma(x)\,\mathrm{d}B_t\right] \\
&= f'(x)\mu(x) + \frac{1}{2}f''(x)(\sigma(x))^2 + \frac{f'(x)\sigma(x)}{\mathrm{d}t}\mathbf{E}\left[\,\mathrm{d}B_t\right]
\end{aligned}$$

$\{B_t\}_{t\geq 0}$ is a standard Brownian motion, thus $\mathbf{E}\left[\,\mathrm{d}B_t\right] = 0$. This indicates that $\mathcal{L}f(x) = f'(x)\mu(x) + \frac{1}{2}f''(x)(\sigma(x))^2$. □

With Lemma 7, we can write the generator of the Itô process explicitly.

Example 2 (The generator of the standard Brownian motion)**.** *The standard Brownian motion satisfies that $\mathrm{d}X_t = \mathrm{d}B_t$. Thus directly applying Lemma 7, we have that*

$$\mathcal{L}f(x) = \frac{1}{2}f''(x)$$

References