STAT0041: Stochastic Calculus

Lecture 13 - Diffusion Process

Lecturer: Weichen Zhao | Fall 2024

Key concepts:

- *Diffusion process;*
- *Markov property of solution of SDE;*
- *Diffusion operator;*
- *Fokker-Planck equation.*

13.1 Diffusion Process

Diffusion is the net movement of anything (for example, atoms, ions, molecules, energy) generally from a region of higher concentration to a region of lower concentration. There is no unique mathematical definition of diffusion, but in probability the key of the diffusion process is that it is a Markov process with continuous sample path.

There have historically been two main methods to studying diffusion processes:

- Analytic method of Kolmogorov:

  Studying the evolution of its macroscopic properties is inscribed by establishing the equations satisfied by transition probability.

- Stochastic calculus method of Itô:

  Studying the trajectory of each particle, that is the SDEs established to describe the trajectory of particles from the microscopic view.

In this lecture we use SDE studying the diffusion process.

Definition 13.1 *A Itô diffusion process is a stochastic process*

$$X_t(\omega) = X(t,\omega)\colon [0,\infty) \times \Omega \to \mathbb{R}^n$$

satisfying a stochastic differential equation of the form

$$dX_t = b(t, X_t)dt + \sigma(t, X_t)dB_t, \quad t \geq s; \quad X_s = x$$

where B_t is m-dimensional Brownian motion and b, σ satisfy the condition of existence and uniqueness of solution:

(1) There exist constant C, such that

$$|b(t,x)|+|\sigma(t,x)|\leqslant C(1+|x|), \quad \forall x\in\mathbb{R}^n, t\in[0,T],$$

(2) There exist constant K, such that

$$|b(t,x)-b(t,y)|+|\sigma(t,x)-\sigma(t,y)|\leqslant K|x-y|, \quad \forall x,y\in\mathbb{R}^n, t\in[0,T].$$

If b, σ only depend on X_t and satisfy exist constant D, such that

$$|b(x)-b(y)|+|\sigma(x)-\sigma(y)|\leq D|x-y|; \quad x,y\in\mathbb{R}^n,$$

we say X_t is a time-homogeneous Itô diffusion process.

Theorem 13.2 *Itô diffusion process is Markov process.*

13.2 Markov semigroup theory

Let us brief review Markov semigroup theory.

Markov semigroup. For Markov process (X_t) with state space $(E,\mathscr{E})$, define:

$$(P_tf)(x)=\mathbb{E}[f(X_t)|X_0=x]=\int f(y)p(t,x,y)dy.$$

where

$$P_t:\mathcal{M}_b(E)\to\mathcal{M}_b(E),\ f\mapsto P_tf$$

$\mathcal{M}_b(E)$ is the set of all bounded measurable function on E. Then P_t has semigroup property:

$$P_{t+s}f=P_t\circ P_sf.$$

We say P_t is a *Markov semigroup.*

Let P_t be a Markov semigroup,

$$\mathcal{L}f:=\lim_{t\downarrow 0}\frac{P_tf-f}{t}.$$

We say $\mathcal{L}$ is the *infinitesimal generator* of P_t.

Adjoint semigroup. The semigroup P_t acts on bounded continuous functions. We can also define the adjoint semigroup P_t^*, which acts on probability measures:

$$(P_t^*\mu)(A) = \int \mathbf{P}(X_t \in A | X_0 = x) d\mu(x) = \int p(t, x, A) d\mu(x).$$

The operators P_t and P_t^* are adjoint, satisfy

$$\int (P_t f)(x) d\mu(x) = \int f(x) d(P_t^*\mu)(x).$$

Denote $\mu_t \triangleq P_t^*\mu$, then μ_t is the law of the Markov process.

Invariant measure. Given a Markov semigroup P_t, a measure π is said to be invariant for P_t if for every bounded positive measurable function $f : E \to \mathbb{R}$ and every $t \geq 0$,

$$\int_E P_t f d\pi = \int_E f d\pi.$$

Equivalently,

$$P_t^*\pi = \pi$$

The Kolmogorov Equations. We have Kolmogorov backward equation

$$\frac{\partial}{\partial t} P_t f = \mathcal{L} P_t f,$$

and Kolmogorov forward equation

$$\frac{\partial}{\partial t} P_t f = P_t \mathcal{L} f,$$

13.3 Diffusion Operators

Definition 13.3 (Diffusion operator) *A Markov diffusion operator $\mathcal{L}$ is a second order differential operator of the form*

$$\begin{aligned}\mathcal{L} f &= \sum_i b_i(x) \frac{\partial}{\partial x_i} f + \frac{1}{2} \sum_{i,j} \Sigma_{ij}(x) \frac{\partial^2}{\partial x_i \partial x_j} f \\ &= b(x) \cdot \nabla f + \frac{1}{2} Tr(\Sigma(x)^T \nabla^2 f)\end{aligned} \tag{13.1}$$

where $\Sigma(x) = (\Sigma_{ij}(x))_{1 \leq i,j \leq n}$ and $b(x) = (b_i(x))_{1 \leq i \leq n}$ are smooth, respectively $n \times n$ symmetric matrix-valued and $\mathbb{R}^n$-valued, functions of x.

Classic diffusion process is constructed from generator, i.e. diffusion operator $\mathcal{L}$. $b(x)$ and $\Sigma(x)$ have clear meaning.

A Markov process X_t in $\mathbb{R}$ with transition function $p(s, x; t, A)$ is called a *diffusion process* if the following conditions are satisfied:

(1) (Continuity). For every x and every $\epsilon > 0$,

$$\lim_{t-s\downarrow 0} \frac{1}{t-s}\, p(s, x; t, \{|X_t - x| > \varepsilon\}) = 0.$$

(2) (Definition of drift coefficient). There exists a function $b(s, x)$ such that for every x and every $\epsilon > 0$,

$$\lim_{t-s\downarrow 0} \frac{1}{t-s} \int_{|y-x|\leq\varepsilon} (y-x)p(s, x; t, \mathrm{d}y) = b(s, x).$$

(3) (Definition of diffusion coefficient). There exists a function $\Sigma(x, s)$ such that for every x and every $\epsilon > 0$,

$$\lim_{t-s\downarrow 0} \frac{1}{t-s} \int_{|y-x|\leq\varepsilon} (y-x)^2 p(s, x; t, \mathrm{d}y) = \Sigma(x, s).$$

Now let us consider generator of Itô diffusion process.

Theorem 13.4 *Let X_t be a time-homogeneous Itô diffusion process*

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

with generator $\mathcal{A}$, then for all twice continuously differentiable function f

$$\mathcal{A}f = \mathcal{L}f$$

where $\mathcal{L}$ defined as (13.1)*, and $\Sigma_{ij}(x) = \sum_{k=1}^{n} \sigma_{ik}(x)\sigma_{kj}(x)$.*

We will use $\mathcal{L}$ denote the generator of Itô diffusion process from now on.

13.4 Fokker-Planck Equation

In this section we focus on the evolution of Law of the diffusion process, which is characterized by Fokker-Planck equation[1].

[1]Fokker–Planck equation in exactly Kolmogorov forward equation, which is more customary in the physics literature.

Theorem 13.5 (Fokker-Planck equation) *Let X_t be a time-homogeneous Itô diffusion process*

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

and $\mu_t(x) = \mu(x,t)$ be the probability density[2] *of X_t at time t, and let $\mu(x)$ be its initial probability density. Then $\mu(x,t)$ solves*

$$\partial_t \mu(x,t) = \mathcal{L}^* \mu(x,t) \quad (t > 0), \quad \mu(x,0) = \mu(x),$$

where operator

$$\begin{aligned} \mathcal{L}^* g &\triangleq -\sum_j \frac{\partial}{\partial x_j}(b_j(x)g) + \frac{1}{2}\sum_{i,j} \frac{\partial^2}{\partial x_i \partial x_j}(\Sigma_{ij}(x)g) \\ &= -\nabla \cdot (b(x)g) + \frac{1}{2}\nabla \cdot \nabla \cdot (\Sigma(x)g) \end{aligned}$$

is adjoint operator of $\mathcal{L}$.

[2] $\mu_t(x)$ is twice continuously differentiable respect to x and t