AM216 Stochastic Differential Equations

Lecture 11
Copyright by Hongyun Wang, UCSC

List of topics in this lecture

- Transition probability density $q(x, t \mid z, s)$ is NOT a probability density in z.
- Derivation of Kolmogorov forward equation, method of test function
- Meaning of backward equation: solution at (z, t) = the average reward at the end time T given starting at position z at time $(T-t)$
- Meaning of forward equation: solution at (x, t) = mass density at time t

---

Recap

Different interpretations of SDE

The Stratonovich interpretation of $dX = b(X,t)dt + \sqrt{a(X,t)}\,dW$ is equivalent to the Ito interpretation of $dX = \left(b(X,t) + \frac{1}{4}a_x(X,t)\right)dt + \sqrt{a(X,t)}\,dW$.

The "correct" interpretation is selected in the modeling process.

Transition probability density (a 4-variable function)

$$q(x, \underbrace{t}_{\substack{\uparrow \\ \text{end time}}} \mid z, \underbrace{s}_{\substack{\uparrow \\ \text{starting time}}}) \equiv \frac{1}{dx}\Pr\left(x < X(t) \le x + dx \,\middle|\, X(s) = z\right), \qquad t > s$$

Backward view (the law of total probability)

We fix (x, t) and view q as a function of (z, s):

$$\underbrace{q(x,t\,|\,z,s)}_{\substack{q(\bullet,s) \\ [s\to t]}} = \int \underbrace{q(x,t\,|\,z+y,s+ds)}_{\substack{q(\bullet,s+ds) \\ [s+ds\to t]}}\ \underbrace{q(z+y,s+ds\,|\,z,s)}_{\substack{\text{density of } dX \\ [s\to s+ds]}}\,dy$$

$$q(\,\cdot\,, s+ds) \quad \longrightarrow \quad q(\,\cdot\,, s)$$

We move the starting time backward from $(s+ds)$ to s.

Forward view (the law of total probability)

We fix (z, s) and view q as a function of (x, t):

$$\underbrace{q(x,t+dt\,|\,z,s)}_{\substack{q(\bullet,t+dt)\\ [s\to t+dt]}} = \int \underbrace{q(x,t+dt\,|\,y,t)}_{\substack{\text{density of } X(t+dt)|X(t)=y\\ [t\to t+dt]}} \; \underbrace{q(y,t\,|\,z,s)}_{\substack{q(\bullet,t)\\ [s\to t]}} \, dy$$

$$q(\cdot\,,t) \quad \longrightarrow \quad q(\cdot\,,t+dt)$$

We move the end time forward from t to $(t+dt)$.

Backward equation, final value problem, converting to IVP.

Before we derive the forward equation, we clarify two issues.

Issue 1: the "correct" interpretation of an SDE is selected in the modeling process.

We compare the two interpretations in the SDE below.

$$dX = \alpha X dW$$

Ito interpretation:

$$X(t+dt) - X(t) = \alpha X(t) dW(t)$$

$$E(X(t+dt) \mid X(t) = x) = x$$

Example 1: Consider a fair game between you and a casino.

Let $X(t)$ = your cash at time t.

- Suppose in each dt, you bet a small random percent of your current cash.

  In each dt, you are equally likely to win or lose that small random percent.

  $$E(X(t+dt) \mid X(t) = x) = x$$

  $$\underbrace{X(t+dt) - X(t) = \alpha X(t) dW(t)}_{\text{Ito}}$$ is appropriate for this situation.

- Suppose in each dt, you bet a small fixed percent of your current cash (for example, you bet $c\sqrt{dt}\,X(t)$ in each dt).

  $$E(X(t+\Delta t) \mid X(t) = x) = x$$

  $$\underbrace{X(t+\Delta t) - X(t) = \alpha X(t) dW(t)}_{\text{Ito}}$$ is appropriate for Δt = many steps of dt.

Stratonovich interpretation

$$X(t+dt) - X(t) = \alpha \frac{X(t) + X(t+dt)}{2} dW(t)$$

Let $Y(t) \equiv \log\big(X(t)\big)$, $dY \equiv Y(t+dt) - Y(t)$.

We examine the evolution of $Y(t)$. We write

$$X(t)=e^{Y(t)}\,,\quad X(t+dt)=e^{Y(t+dt)}=e^{Y(t)+dY}$$

Substitute into the SDE for X, divide by $e^{Y(t)+dY/2}$ and expand in dY

$$e^{Y(t)+dY}-e^{Y(t)}=\frac{\alpha}{2}\left(e^{Y(t)}+e^{Y(t)+dY}\right)dW(t)$$

$$e^{dY/2}-e^{-dY/2}=\frac{\alpha}{2}\left(e^{dY/2}+e^{-dY/2}\right)dW(t)$$

$$dY+O\left((dY)^3\right)=\alpha\left[1+O\left((dY)^2\right)\right]dW(t)$$

<u>Remark:</u> dividing by $e^{Y(t)+dY/2}$ made it simple!

Use $dY=\log(X(t)+dX)-\log X(t)\approx\dfrac{dX}{X(t)}\approx dW=O(\sqrt{dt})$ and neglect o(dt) terms.

$$dY=\alpha\,dW(t)$$

$$\log\left(X(t+dt)\right)-\log\left(X(t)\right)=\alpha dW(t)\quad\Rightarrow\quad X(t+dt)=X(t)e^{\alpha dW(t)}$$

$$E\left(\log X(t+dt)\middle|X(t)=x\right)=\log x$$

$$E\left(X(t+dt)\middle|X(t)=x\right)=E\left(x\,e^{\alpha dW(t)}\right)=x\,E\left(1+\alpha dW+\tfrac{1}{2}\alpha^2(dW)^2+o(dt)\right)$$
$$=x\left(1+\tfrac{1}{2}\alpha^2dt+o(dt)\right)>x$$

<u>Example 2:</u>

Consider a <u>game</u> between you and a "casino" (<u>the stock market</u>).

Let $X(t)$ = your net worth at time t.

Suppose in each dt, you net worth is equally likely to be multiplied or divided by a <u>factor close to 1</u> (for example, a factor of $\left(1+c\sqrt{dt}\right)$)

$$E\left(\log X(t+dt)\middle|X(t)=x\right)=\log x$$

$$\underbrace{X(t+dt)-X(t)=\alpha\frac{X(t)+X(t+dt)}{2}dW(t)}_{\text{Stratonovich}}$$ is appropriate for this situation.

$$E\left(X(t+dt)\middle|X(t)=x\right)=x\frac{1}{2}\left(1+c\sqrt{dt}+\frac{1}{1+c\sqrt{dt}}\right)=x\left(1+\frac{1}{2}c^2dt+o(dt)\right)>x$$

$$\underbrace{X(t+dt)-X(t)=\alpha X(t)dW(t)}_{\text{Ito}}$$ is not appropriate for this situation.

<u>The two interpretations are related to each other.</u>

Stratonovich of $dX = \mu X dt + \alpha X dW(t)$

is equivalent to Ito of the modified SDE, $dX = \left(\mu X + \frac{1}{2}\alpha^2 X\right)dt + \alpha X dW(t)$.

Now back to the discussion of transition probability density.

Issue 2: In general, $q(x, t \mid z, s)$ is NOT a probability density in z.

Example: Ornstein-Uhlenbeck process

$$dY = -\beta Y dt + \sqrt{\gamma^2}\, dW$$

Recall that previously we derived

$$\left(Y(t) \middle| Y(0) = y_0\right) \sim N\left(\mathrm{e}^{-\beta t} y_0,\ \frac{\gamma^2}{2\beta}(1 - e^{-2\beta t})\right)$$

For simplicity, we set $\beta = 1$ and $\gamma^2/(2\beta) = 1$. Apply it to $[s, t]$ with $t > s$.

$$\left(Y(t) \middle| Y(s) = z\right) \sim N\left(\mathrm{e}^{-(t-s)} z,\ (1 - \mathrm{e}^{-2(t-s)})\right)$$

The transition probability density of Y is

$$q(x, t \mid z, s) = \frac{1}{\sqrt{2\pi(1 - \mathrm{e}^{-2(t-s)})}} \exp\left(\frac{-(x - \mathrm{e}^{-(t-s)} z)^2}{2(1 - \mathrm{e}^{-2(t-s)})}\right)$$

As a function of x, it is a probability density.

We examine it as a function of z. For simplicity, we set $s = 0$, $t = 1$ and $x = 0$.

$$q(0, 1 \mid z, 0) = \frac{1}{\sqrt{2\pi(1 - \mathrm{e}^{-2})}} \exp\left(\frac{-(0 - \mathrm{e}^{-1} z)^2}{2(1 - \mathrm{e}^{-2})}\right) = \frac{\mathrm{e}^1}{\sqrt{2\pi(\mathrm{e}^2 - 1)}} \exp\left(\frac{-z^2}{2(\mathrm{e}^2 - 1)}\right)$$

$$= \mathrm{e}^1 \cdot \rho_{N\left(0,\, (\mathrm{e}^2 - 1)\right)}(z) \qquad \text{not a probability density}$$

Key observations:

- We should not expect $\int q(x, t \mid z, s) dz = 1$
- We should not expect $\int q(x, t \mid z, s) dz$ to be conserved with respect to s.

Derivation of the forward equation for SDE $dX = b(X, t)dt + \sqrt{a(X, t)}\, dW$

We fix (z, s) and view q as a function of (x, t): $\qquad q(x, t) \equiv q(x, t \mid z, s)$

Forward view:

$$\underbrace{q(x,t+dt\,|\,z,s)}_{q(\bullet,t+dt)} = \int \underbrace{q(x,t+dt\,|\,y,t)}_{\substack{\text{density of}\\ X(t+dt)|X(t)=y}}\ \underbrace{q(y,t\,|\,z,s)}_{q(\bullet,t)}\,dy$$

Key for the derivation:

As $dt \to 0$, $q(x,t+dt\,|\,y,t)$ is significant only for small $|y-x|$. As a result, the integral is dominated by contribution from small $|y-x|$.

The old approach of expanding $q(y, t)$ around $y = x$ won't work!

$$q(y,t) = q(x+(y-x),t) = \cdots + q_x(x,t\,|\,z,s)(y-x) + \cdots$$

$$\int q(x,t+dt\,|\,y,t)q(y,t)dy = \cdots + \int \underbrace{q\big(x,t+dt\,|\,y,t\big)}_{\substack{\text{This is NOT}\\ \text{a density in } y!}}\ \underbrace{q_x(x,t)}_{\substack{\text{This is fine. It is}\\ \text{independent of } y}}\ (y-x)dy + \cdots$$

- Integrating a density leads to moments.

$$\int (x-y)\underbrace{q(x,t+dt\,|\,y,t)}_{\text{This is a density in } x.}\,dx = E(dX)$$

- Integrating a non-density leads to nowhere.

$$\int (y-x)\underbrace{q(x,t+dt\,|\,y,t)}_{\substack{\text{This is NOT}\\ \text{a density in } y!}}\,dy = \text{unknown}$$

Strategy:

We multiply it by a test function $h(x)$ and then integrate with respect to x.

Implementation

Let $h(x)$ be a smooth function with compact support.

Definition:

We say function $h(x)$ has compact support if there exists M such that

$$h(x) = 0 \quad \text{for} \quad |x| > M$$

We multiply both sides of the master equation by $h(x)$ and integrate with respect to x.

$$\text{LHS} = \int q(x,t+dt)h(x)dx = \int\big[q(x,t) + q_t dt + o(dt)\big]h(x)dx$$

$$\text{RHS} = \int\Big[\int q(y,t)q(x,t+dt\,|\,y,t)dy\Big]h(x)dx$$

Changing the order of integration leads to

$$\text{RHS} = \int q(y,t)\left[\int \underbrace{q(x,t+dt\,|\,y,t)}_{\text{This is a density in } x.}\,h(x)dx\right]dy \tag{E01}$$

The inner integral is dominated by contribution from small $|x-y|$.

We expand $h(x)$ around $x = y$.

$$h(x)=h(y+(x-y))=h(y)+h_y(y)(x-y)+\frac{h_{yy}(y)}{2}(x-y)^2+O\left((x-y)^3\right)$$

In the inner integral, these expansion terms lead to moments of dX.

Recall the moments of dX for the SDE $dX=b(X,t)dt+\sqrt{a(X,t)}\,dW$ (Ito).

$$E\left((dX)^0\right)=\int q(x,t+dt\,|\,y,t)dx=1$$

$$E\left((dX)^1\right)=\int q(x,t+dt\,|\,y,t)(x-y)dx=b(y,t)dt+o(dt)$$

$$E\left((dX)^2\right)=\int q(x,t+dt\,|\,y,t)(x-y)^2dx=a(y,t)dt+o(dt)$$

$$E\left((dX)^n\right)=\int q(x,t+dt\,|\,y,t)(x-y)^n dx=o(dt)\,,\qquad n\geq 3$$

The inner integral becomes

$$\int q(x,t+dt\,|\,y,t)h(x)dx=h(y)+h_y(y)b(y,t)dt+\frac{h_{yy}(y)}{2}a(y,t)dt+o\left(dt\right)$$

- Substituting this result into the outer integral in (E01),
- integrating by parts, and using the compactness of $h(y)$, we get

$$\text{RHS}=\int\left[q-\left(b(y,t)q\right)_y dt+\frac{1}{2}\left(a(y,t)q\right)_{yy}dt\right]h(y)dy+o(dt),\qquad q\equiv q(y,t)$$

- Renaming variable y as x, we write it as

$$\text{RHS}=\int\left[q-\left(b(x,t)q\right)_x dt+\frac{1}{2}\left(a(x,t)q\right)_{xx}dt\right]h(x)dx+o(dt),\qquad q\equiv q(x,t)$$

- Subtracting $\int q(x,t)h(x)dx$ from both LHS and RHS,
- dividing by dt, and taking the limit as $dt\to 0$, we arrive at

$$\text{LHS}=\int q_t h(x)dx$$

$$\text{RSH}=\int\left[-\left(b(x,t)q\right)_x+\frac{1}{2}\left(a(x,t)q\right)_{xx}\right]h(x)dx,\qquad q\equiv q(x,t)$$

Since LHS = RHS for all test function $h(x)$, we conclude

$$q_t=-\left(b(x,t)q\right)_x+\frac{1}{2}\left(a(x,t)q\right)_{xx}$$

This is called the Fokker-Planck equation or the Kolmogorov forward equation.

Conservation form:

The forward equation has the conservation form

$$q_t = -\frac{\partial}{\partial x} J(x,t)$$

where $J(x,t) \equiv b(x,t)q - \frac{1}{2}\big(a(x,t)q\big)_x$ is the probability flux

Terminology: flux ≡ flow per unit time

Remarks:

- Solution of $q_t = -\frac{\partial}{\partial x} J(x,t)$ is conserved:

$$\int_a^b q(x,t_2)dx - \int_a^b q(x,t_1)dx = \underbrace{\int_{t_1}^{t_2} J(a,t)dt}_{\text{In-flow}} - \underbrace{\int_{t_1}^{t_2} J(b,t)dt}_{\text{Out-flow}}$$

  Change in $\int_a^b q(x,t)dx$ is attributed to in-flow at $x = a$ and out-flow at $x = b$.

- In contrast, the backward equation is not in the conservation form.

$$q_s = -b(z,s)q_z - \frac{1}{2}a(z,s)q_{zz}$$

  In general, solution of the backward equation is not conserved.

The initial value problem (IVP) for $q(x, t) \equiv q(x, t \mid z, 0)$

$$\begin{cases} q_t = -\big(b(x,t)q\big)_x + \frac{1}{2}\big(a(x,t)q\big)_{xx} \\ q(x,t \mid z,0)\big|_{t=0} = \delta(x-z) \end{cases}$$

We solve it forward from $t = 0$ to $t = T$.

Autonomous SDEs:

$$dX = b(X)dt + \sqrt{a(X)}\,dW, \qquad b(x,t) = b(x), \quad a(x,t) = a(x)$$

- There is no explicit dependence on time.
- If we shift in time, the evolution remains the same.

The IVP of backward equation has a simple form in the autonomous case.

Backward equation in the autonomous case:

We shift in time by $(T-\tau)$ to write

$\phi(z, \tau) \equiv q(x, T \mid z, T-\tau) = q(x, \tau \mid z, 0)$ where x and T are fixed.

The IVP for $\phi(z, \tau)$ is

$$\begin{cases} \phi_\tau = \beta(z,\tau)\phi_z + \dfrac{1}{2}\alpha(z,\tau)\phi_{zz}\,, \qquad \beta(z,\tau) \equiv b(z,T-\tau)\,, \quad \alpha(z,\tau) \equiv a(z,T-\tau) \\ \phi(z,\tau)\big|_{t=0} = \delta(z-x) \end{cases}$$

In the autonomous case, $\beta(z, \tau) = b(z)$, $\alpha(z, \tau) = a(z)$. For simplicity, we change back to (deceptively) simple notations:

$$\tau \to t, \qquad \phi(z, \tau) \to q(z, t).$$

The IVP for $q(z, t) \equiv q(x, T \mid z, T - t) = q(x, t \mid z, 0)$ is

$$\begin{cases} q_t = b(z)q_z + \dfrac{1}{2}a(z)q_{zz} \\ q(z,t)\big|_{t=0} = \delta(z-x) \end{cases} \qquad \text{where } x \text{ is a parameter.}$$

Remark: In applications, end time T is fixed and t in $q(z, t)$ refers to the time until the end time, corresponding to real time $(T-t)$.

Forward equation in the autonomous case:

The IVP for $q(x, t) \equiv q(x, t \mid z, 0)$ is

$$\begin{cases} q_t = -\big(b(x)q\big)_x + \dfrac{1}{2}\big(a(x)q\big)_{xx} \\ q(x,t)\big|_{t=0} = \delta(x-z) \end{cases} \qquad \text{where } z \text{ is a parameter.}$$

Meaning of the backward equation with a *general initial condition*

We consider the autonomous SDE $dX = b(X)dt + \sqrt{a(X)}\,dW$.

$$\begin{cases} u_t = b(z)u_z + \dfrac{1}{2}a(z)u_{zz} \\ u(z,t)\big|_{t=0} = u_0(z) \end{cases} \qquad \text{(BE\_IVP1)}$$

Recall two examples we studied.

- For the transition PD $q(z, t) \equiv q(x, T \mid z, T - t) = q(x, t \mid z, 0)$

$$q(z,t)\big|_{t=0} = \delta(z-x)$$

- For the probability of winning bet $X(T) \geq x_C$ given $X(T-t) = z$

$$u(z,t) = \Pr\big(X(T) \geq x_c \,\big|\, X(T-t) = z\big)$$

$$u(z,t)\Big|_{t=0} = \begin{cases} 1\,, & z \ge x_C \\ 0\,, & z < x_C \end{cases}$$

Here we look at a general initial condition: $u(z, t)|_{t=0} = u_0(z)$.

It is straightforward to verify that the solution of (BE_IVP1) is

$$u(z,t) = \int q(x,T\,|\,z,T-t)u_0(x)dx \qquad \text{(B01)}$$

<u>Observations</u>:

- $q(x, T \mid z, T\text{–}t)$ is the transition probability density.
- Real time T is a future time, for example, the expiration date of an option.
- Variable t in the backward equation is the time until the end time. Variable t corresponds to *real time* (T–t).

<u>Meaning of solution $u(z, t)$</u>

Suppose the reward is determined <u>at *real time T*</u>, based on $X(T)$.

Let $u_0(x)$ be the reward function, which maps $X(T)$ to reward:

the amount of reward = $u_0(X(T))$

<u>Example</u>:

Let $X(t)$ = the market price of a stock at time t.

Consider a "call" option to buy the stock at price x_C at time T.

<u>Terminology:</u>

A call option = the right (not obligation) to buy a certain number of shares of the stock at a specified price at a preset time (expiration date).

A put option = the right (not obligation) to sell a certain number of shares of the stock at a specified price at a preset time (expiration date).

The amount of reward for owning the call option depends on $X(T)$. It is

$$u_0\big(X(T)\big) = \begin{cases} X(T) - x_C\,, & X(T) > x_C \\ 0\,, & X(T) \le x_C \end{cases}$$

Suppose X starts at position z at <u>*real time* (T–t)</u>. The conditional distribution of $X(T)$ given $X(T$–$t)$ = z is described by the transition PD

$q(x, T \mid z, T\text{–}t)$

The conditional expected amount of reward given $X(T$–$t)$ = z is

$$E\Big(u_0\big(X(T)\big)\Big|X(T-t)=z\Big) = \int \underbrace{q(x,T\,|\,z,T-t)}_{\substack{\text{transition density of} \\ X(T)=x|X(T-t)=z}}\ \underbrace{u_0(x)}_{\substack{\text{reward} \\ \text{for } X(T)=x}}\ dx \qquad \text{(B02)}$$

This is exactly the same as the solution $u(z, t)$ given in (B01).

Summary (meaning of the backward equation)

Suppose the reward is determined at *real time T*, as $u_0(X(T))$.

The solution of the backward equation, $u(z, t)$, is the expected amount of reward given $X(T-t) = z$.

$$\underbrace{u(z,t)}_{\substack{\text{solution of the}\\ \text{backward equation}}} = \underbrace{E\Big(u_0\big(X(T)\big)\Big|X(T-t)=z\Big)}_{\substack{\text{expected amount of reward}\\ \text{given } X(T-t)=z}}$$

The backward equation describes the backward time evolution of the expected amount of reward. The end time is fixed at T. In the backward time evolution, the start time is gradually moved backward from T to $(T-t)$.

In general, the expected amount of reward $q(z, t)$ is not conserved.

$$\int q(z,t_1)dz \neq \int q(z,t_2)dz$$

This is related to that the backward equation is NOT in the conservation form.

Meaning of the forward equation with a *general initial condition*

We consider the autonomous SDE $dX = b(X)dt + \sqrt{a(X)}\,dW$

$$\begin{cases} p_t = -\big(b(x)p\big)_x + \dfrac{1}{2}\big(a(x)p\big)_{xx} \\ p(x,t)\big|_{t=0} = p_0(x) \end{cases} \qquad \text{(FE\_IVP1)}$$

It is straightforward to verify that the solution of (FE_IVP1) is

$$p(x,t) = \int \underbrace{q(x,t\,|\,z,0)}_{\text{start time is 0}}\, p_0(z)\,dz \qquad \text{(F01)}$$

Observations:

- $q(x, t \mid z, 0)$ is the transition probability density.
- Variable t in the forward equation is the time elapsed since the start time.

Meaning of solution $p(x, t)$

Consider a set of X. For example, a set of 7 particles, $\{X_j(t), j = 1, 2, \ldots, 7\}$.

All averages are based on an ensemble of the set.

Ensemble = $\{\ \{X_j(t, \omega), j = 1, 2, \ldots, 7\},\ \omega \in \Omega\ \}$

a collection of an infinite number of *independent* copies of the set.

Mass density at position x at time t is

$$\rho(x,t)=\frac{1}{dx}E_{\omega}\Big(\#\text{ of X}_j(t,\omega)\text{'s in }[x,x+dx]\Big)=\sum_{j=1}\rho(X_j(t)=x)$$

Notation: $\rho(X{=}x) \equiv \rho_X(x)$.

Let $p_0(x)$ be the mass density at time 0.

The mass density at time t is given by the law of total probability.

$$\sum_{j=1}\rho(X_j(t)=x)=\sum_{j=1}\int\underbrace{\rho(X_j(t)=x\,\big|\,X_j(0)=z)}_{q(x,t|z,0),\text{ independent of }j}\rho(X_j(0)=z)dz$$

$$=\int q(x,t\,\big|\,z,0)\underbrace{\sum\nolimits_{j=1}\rho(X_j(0)=z)}_{p_0(z)}dz=\int q(x,t\,\big|\,z,0)p_0(z)dz \tag{F02}$$

This is exactly the same as the solution $p(x, t)$ given in (F01).

Summary:

We consider a set of X because the mass density is more general.

Suppose $p_0(x)$ is the mass density of X at time 0.

The solution of the forward equation, $p(x, t)$ = the mass density of X at time t.

The forward equation describes the forward time evolution of mass density.

The mass density $p(x, t)$ is conserved.

$$\int_a^b p(x,t_2)dx-\int_a^b p(x,t_1)dx=\underbrace{\int_{t_1}^{t_2}J(a,t)dt}_{\text{In-flow}}-\underbrace{\int_{t_1}^{t_2}J(b,t)dt}_{\text{Out-flow}}$$

This is related to that the forward equation is in the conservation form.

A tricky issue: When the number of particle in the set is very large, one copy is enough for describing the behavior of the ensemble and thus is often called an ensemble. In that case, mass density is also called ensemble density.

Example: an ensemble of 3×10^{16} air molecules in a volume of 1 $(\text{mm})^3$.