Lecture 7: Stochastic integrals and stochastic differential equations

Eric Vanden-Eijnden

Combining equations (1) and (2) from Lecture 6, one sees that W_t^N satisfies the recurrence relation

$$W_{t_n}^N = W_{t_n}^N + \xi_{n+1}\sqrt{\Delta t}, \qquad W_0^N = 0. \tag{1}$$

where $t_n = n/N$, $\Delta t = 1/N$ and $\{\xi_n\}_{n\in\mathbb{N}}$ are i.i.d. random variables taking values ± 1 with probability $\frac{1}{2}$ as before. A natural generalization of this relation is

$$X_{t_{n+1}}^N = X_{t_n}^N + b(X_{t_n}^N, t_n)\Delta t + \sigma(X_{t_n}^N, t_n)\xi_{n+1}\sqrt{\Delta t}, \qquad X_0 = x \tag{2}$$

If the last term were absent, this would be the forward Euler scheme for the ordinary differential equation (ODE) $\dot{X}_t = b(X_t, t)$. If $b(x,t)$ and $\sigma(x,t)$ meet appropriate regularity requirements, it can be shown that X_t^N converges to a stochastic process X_t as $N \to \infty$ (i.e. as $\Delta t \to 0$ with $n\Delta t \to t$). The limiting equation for X_t is denoted as the *stochastic differential equation* (SDE)

$$dX_t = b(X_t, t)dt + \sigma(X_t, t)dW_t, \qquad X_0 = x, \tag{3}$$

as a remainder that the last term in (2) divided by Δt does not have a standard function as limit. The notation dW_t comes from (1) since this equation can be written as $W_{t_{n+1}}^N - W_{t_n}^N = \xi_{n+1}\sqrt{\Delta t}$. We note that the convergence of X_t^N to X_t holds provided only that the ξ_n's are i.i.d. random variables with mean zero, $\mathbb{E}\xi_n = 0$, and variance one, $\mathbb{E}\xi_n^2 = 1$. The standard choice in numerical schemes is to take $\xi_n = N(0,1)$, in which case

$$\sqrt{\Delta t}\,\xi_{n+1} \stackrel{d}{=} W_{t_{n+1}} - W_{t_n}.$$

In the discussion below, however, we will stick to the choice where $\{\xi_n\}_{n\in\mathbb{N}}$ are i.i.d. random variables taking values ± 1 with probability $\frac{1}{2}$ since it facilitates the calculations.

Next, we study the properties of X_t solution of (3) and introduce some nonstandard calculus due to Itô to manipulate this solution.

1 Itô isometry and Itô formula

Consider the recurrence relation

$$X_{t_{n+1}}^N = X_{t_n}^N + f(W_{t_n}^N)\xi_{n+1}\sqrt{\Delta t}, \qquad X_0^N = 0.$$

Let us investigate the properties of the limit of $X^N_{n\Delta t}$ as $N \to \infty$, assuming that this limit exists. The limiting form of the recurrence relation above is traditionally denoted as

$$dX_t = f(W_t, t)dW_t, \qquad X_0 = 0,$$

which can also be expressed as the *stochastic integral*

$$X_t = \int_0^t f(W_s, s)dW_s.$$

Stochastic integral have special properties called the *Itô isometry*

$$\begin{aligned}
&\mathbb{E}\int_0^t f(W_s, s)dW_s = 0,\\
&\mathbb{E}\Big(\int_0^t f(W_s, s)dW_s\Big)^2 = \int_0^t \mathbb{E}f^2(W_s, s)ds.
\end{aligned}$$

The first of these identity is often written and used in differential form

$$\mathbb{E}f(W_s, s)dW_s = 0.$$

The Itô isometry is easy to demonstrate. The first identity is implied by

$$\begin{aligned}
\mathbb{E}X^N_{t_n} &= \mathbb{E}\sum_{m=0}^{n-1} f(W^N_{t_m}, t_m)\xi_{m+1}\sqrt{\Delta t}\\
&= \sum_{m=0}^{n-1} \mathbb{E}f(W^N_{t_m}, t_m)\mathbb{E}\xi_{m+1}\sqrt{\Delta t} = 0,
\end{aligned}$$

where we used the independence of the ξ_m's and $\mathbb{E}\xi_m = 0$. The second identity is implied by

$$\begin{aligned}
\mathbb{E}(X^N_{t_n})^2 &= \mathbb{E}\sum_{m,p=0}^{n} f(W^N_{t_m}, t_m)f(\bar{W}^N_{t_p}, t_p)\xi_{m+1}\xi_{p+1}\Delta t\\
&= \sum_{m=0}^{n} \mathbb{E}f^2(W^N_{t_m}, t_m)\Delta t,
\end{aligned}$$

where we use the fact that ξ_m and ξ_p are independent unless $m = p$, and $\xi_m^2 = 1$ by definition.

Going back to (3), a very important formula to manipulate the solution of this equation is *Itô formula* which states the following. Assume that X_t is the solution of (3) and let f be a smooth function. Then $g(X_t)$ satisfies the SDE

$$\begin{aligned}
dg(X_t) &= g'(X_t)dX_t + \tfrac{1}{2}g''(X_t)\sigma^2(X_t, t)dt\\
&= \big(g'(X_t)b(X_t, t) + \tfrac{1}{2}g''(X_t)\sigma^2(X_t, t)\big)\,dt + g'(X_t)\sigma(X_t, t)dW_t.
\end{aligned}$$

If g depends explicitly on t, then an additional term $\partial g/\partial t dt$ is present at the right hand-side. Itô formula is the analog of the chain rule in ordinary differential calculus. However ordinary chain rule would give

$$dg(X_t) = g'(X_t)dX_t.$$

Here because of the non-differentiability of X_t, we have the additional term that depends on $g''(x)$.

The proof of Itô formula can be outlined as follows. We Taylor expand $g(X_{t_{n+1}}^N) - g(X_{t_n}^N)$ using the recurrence relation (2) for $X_{t_n}^N$ and keep terms up to $O(\Delta t)$:

$$
\begin{aligned}
&g(X_{t_{n+1}}^N) - g(X_{t_n}^N) \\
&\quad = g'(X_{t_n}^N)(X_{t_{n+1}}^N - X_{t_n}^N) + \tfrac{1}{2} g''(X_{t_n}^N)(X_{t_{n+1}}^N - X_{t_n}^N)^2 + \cdots \\
&\quad = g'(X_{t_n}^N)(X_{t_{n+1}}^N - X_{t_n}^N) \\
&\qquad\quad + \tfrac{1}{2} g''(X_{t_n}^N)\Big(b(X_{t_n}^N, t_n)\Delta t + \sigma(X_{t_n}^N, t_n)\xi_{n+1}\sqrt{\Delta t}\Big)^2 + O(\Delta t^{3/2}) \\
&\quad = g'(X_{t_n}^N)(X_{t_{n+1}}^N - X_{t_n}^N) + \tfrac{1}{2} g''(X_{t_n}^N)\sigma^2(X_{t_n}^N, t_n)\xi_{n+1}^2 \Delta t + O(\Delta t^{3/2}) \\
&\quad = g'(X_{t_n}^N)(X_{t_{n+1}}^N - X_{t_n}^N) + \tfrac{1}{2} g''(X_{t_n}^N)\sigma^2(X_{t_n}^N, t_n)\Delta t + O(\Delta t^{3/2}),
\end{aligned}
$$

where in the last equality we used $\xi_{n+1}^2 = 1$. The Itô formula follows in the limit as $\Delta t \to 0$.

2 Examples

The Itô isometry and the Itô formula are the backbone of the *Itô calculus* which we now use to compute some stochastic integrals and solve some SDEs. As an example of stochastic integral, consider

$$\int_0^t W_s dW_s.$$

Taking $f(x) = x^2$ in Itô formula gives

$$\tfrac{1}{2} dW_t^2 = W_t dW_t + \tfrac{1}{2} dt.$$

Therefore

$$\int_0^t W_s dW_s = \tfrac{1}{2} W_t^2 - \tfrac{1}{2} t.$$

Notice that the second term at the right hand-side would be absent by the rules of standard calculus. Yet, this term must be present for consistency, since the expectation of the left hand-side is

$$\mathbb{E}\int_0^t W_s dW_s = 0,$$

using the first Itô isometry, and the expectation of the right hand-side is zero only with the term $\frac{1}{2}t$ included since $\frac{1}{2}\mathbb{E}W_t^2 = \frac{1}{2}t$.

As a first example of SDE, consider

$$dX_t = -\gamma X_t dt + \sigma dW_t, \qquad X_0 = x$$

This is the *Ornstein-Uhlenbeck process*. Using Itô formula with $f(x,t) = e^{\gamma t} x$, we get (this is Duhammel principle)

$$d\big(e^{\gamma t} X_t\big) = \gamma e^{\gamma t} X_t dt + e^{\gamma t} dX_t = \sigma e^{\gamma t} dW_t.$$

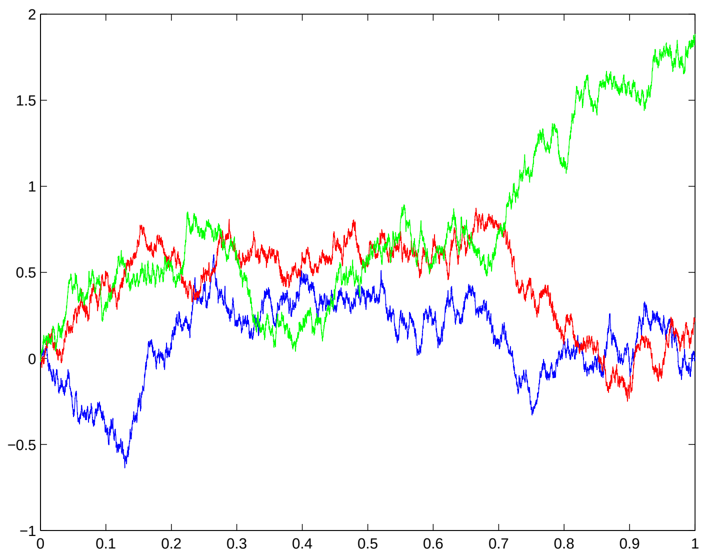

Figure 1: Three realizations of the Ornstein-Uhlenbeck process with $X_0 = 0$ and $\gamma = \sigma = 1$.

Integrating gives

$$X_t = e^{-\gamma t}x + \sigma \int_0^t e^{-\gamma(t-s)} dW_s.$$

This process is Gaussian being a linear combination of the Gaussian process W_t. Its mean and variance are (using the Itô isometry)

$$\mathbb{E}X_t = e^{-\gamma t}x$$
$$\mathbb{E}(X_t - \mathbb{E}X_t)^2 = \sigma^2 \int_0^t (e^{-\gamma(t-s)})^2 ds = \frac{\sigma^2}{2\gamma}(1 - e^{-2\gamma t}).$$

Thus when $\gamma > 0$

$$X_t \xrightarrow{d} N\left(0, \frac{\sigma^2}{2\gamma}\right),$$

as $t \to \infty$.

As a second example of SDE, consider the so-called *geometric Brownian motion*

$$dY_t = Y_t dt + \alpha Y_t dW_t, \qquad Y_0 = y.$$

This process has some application in mathematical finance. Itô's formula with $f(x) = \log x$ gives

$$d \log Y_t = \frac{1}{Y_t}(Y_t dt + \alpha Y_t dW_t) - \frac{1}{2Y_t^2}\alpha^2 Y_t^2 dt.$$

Integrating we get

$$Y_t = y e^{t - \frac{1}{2}\alpha^2 t + \alpha W_t}.$$

Note that by the rules of standard calculus, we would have obtained the wrong answer

$$Y_t = y e^{t + \alpha W_t} \qquad \text{(wrong!)}$$

Indeed the term $-\frac{1}{2}\alpha^2 t$ in the exponential is important for consistency since taking the expectation of the SDE for Y_t using the first Itô isometry gives

$$d\mathbb{E}Y_t = \mathbb{E}Y_t dt,$$

and hence

$$\mathbb{E}Y_t = ye^t.$$

The solution above is consistent with this since

$$\mathbb{E}e^{\alpha W_t} = e^{\frac{1}{2}\alpha^2 t}.$$

3 Generalization in multi-dimension

The definition of Itô integrals and SDE's can be extended to multi-dimension in a straightforward fashion. The SDE

$$dX_t^j = b_j(X_t, t)dt + \sum_{k=1}^{K} \sigma_{jk}(X_t, t)dW_t^k, \qquad j = 1, \dots, J$$

where $\{W_t^k\}_{k=1}^K$ are independent Wiener processes, defines a vector-valued stochastic process $X_t = (X_t^1, \dots, X_t^J)$. The only point worth noting is the Itô formula, which in multi-dimension reads:

$$df(X_t) = \sum_{j=1}^{J} \frac{\partial f(X_t)}{\partial x_j} dX_t^j + \frac{1}{2} \sum_{j,j'=1}^{J} \frac{\partial^2 f(X_t)}{\partial x_j \partial x_{j'}} \Big(\sum_{k=1}^{K} \sigma_{jk}(X_t, t)\sigma_{kj'}(X_t, t) \Big) dt$$

4 Forward and backward Kolmogorov equations

Consider the stochastic ODE

$$dX_t = b(X_t)dt + \sigma(X_t)dW_t, \qquad X_0 = y.$$

Define the transition probability density $\rho(x, t|y)$ via

$$\int_{x_1}^{x_2} \rho(x, t|y)dx = \mathbb{P}\{X_{t+s} \in [x_1, x_2)|X_s = y\}.$$

($\rho(x, t|y)$ does not depends on s because $b(x)$ and $\sigma(x)$ are time-independent.) The transition probability density is an essential object because the process X_t is *Markov*, in other words: for any $t, s \geq 0$

$$\mathbb{P}(X_{t+s} \in B[x_1, x_2)|\{X_{s'}\}_{0\leq s' \leq s}) = \mathbb{P}(X_{t+s} \in B[x_1, x_2)|\{X_s\}),$$

i.e.the future behavior of X_t given what has happened up to time s depends only on what X_s was. We will derive equation for ρ. Let f be an arbitrary smooth function. Using Itô formula, we have

$$f(X_t) - f(y) = \int_0^t f'(X_s)dX_s + \frac{1}{2} \int_0^t f''(X_s)a(X_s)ds,$$

where $a(x) = \sigma^2(x)$. Taking expectation on both sides, we get

$$\mathbb{E}f(X_t) - f(y) = \mathbb{E}\int_0^t f'(X_s)b(X_s)ds + \tfrac{1}{2}\mathbb{E}\int_0^t f''(X_s)(X_s)ds.$$

or equivalently using ρ

$$\begin{aligned}&\int_{\mathbb{R}} f(x)\rho(x,t|y)dx - f(y)\\ &= \int_0^t \int_{\mathbb{R}} f'(x)b(x)\rho(x,s|y)dxds + \tfrac{1}{2}\int_0^t \int_{\mathbb{R}} f''(x)a(x)\rho(x,s|y)dxds.\end{aligned}$$

Since this holds for all smooth f, we obtain

$$\frac{\partial \rho}{\partial t} = -\frac{\partial}{\partial x}(b(x)\rho) + \tfrac{1}{2}\frac{\partial^2}{\partial x^2}(a(x)\rho) \tag{4}$$

with the initial condition $\lim_{t\to 0}\rho(x,t|y) = \delta(x-y)$. This is the *forward Kolmogorov equation* for ρ in terms of the variables (x,t). It is also called the *Fokker-Planck equation.*

Equivalently, an equation for ρ in terms of the variables (y,t) can be derived. The Markov property implies that

$$\rho(x,t+s|y) = \int_{\mathbb{R}} \rho(x,t|z)\rho(z,s|y)dz.$$

Hence

$$\begin{aligned}\rho(x,t+\Delta t|y) - \rho(x,t|y) &= \int_{\mathbb{R}} \rho(x,t|z)\rho(z,\Delta t|y)dz - \rho(x,t|y)\\ &= \int_{\mathbb{R}} \rho(x,t|z)\big(\rho(z,\Delta t|y) - \delta(z-y)\big)dz.\end{aligned}$$

Dividing both side by Δt and taking the limit as $\Delta t \to 0$ using the forward Kolmogorov equation one obtains

$$\frac{\partial \rho}{\partial t} = \int_{\mathbb{R}} \rho(x,t|z)\Big(-\frac{\partial}{\partial z}(b(z)\delta(z-y)) + \tfrac{1}{2}\frac{\partial^2}{\partial z^2}(a(z)\delta(z-y))\Big)dz,$$

which by integration by parts gives

$$\frac{\partial \rho}{\partial t} = b(y)\frac{\partial \rho}{\partial y} + \tfrac{1}{2}a(y)\frac{\partial^2 \rho}{\partial y^2}. \tag{5}$$

This is the *backward Kolmogorov equation* for ρ in terms of the variables (y,t). The operator

$$L = b(y)\frac{\partial}{\partial y} + \tfrac{1}{2}a(y)\frac{\partial^2}{\partial y^2},$$

is called the *infinitesimal generator* of the process. The coefficient b and a can be expressed as

$$b(y) = \lim_{t\to 0}\frac{1}{t}(\mathbb{E}_y X_t - y), \qquad a(y) = \lim_{t\to 0}\frac{1}{t}\mathbb{E}_y(X_t - y)^2,$$

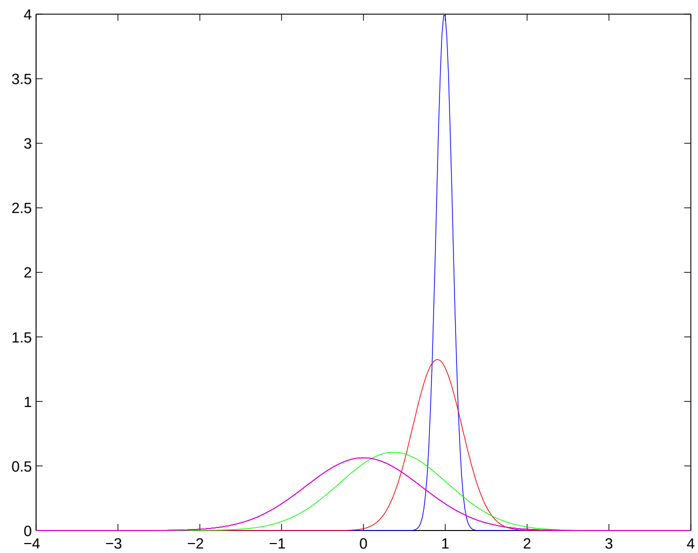

Figure 2: Snapshots of the density of the Ornstein-Uhlenbeck process at time $t = 0.01$ (blue), $t = 0.1$ (red), $t = 1$ (green), and $t = 10$ (magenta). Here $X_0 = y = 1$ and $\gamma = \sigma = 1$. The last snapshot at $t = 10$ is very close to the equilibrium density.

where $\mathbb{E}_y$ denotes expectation conditional on $X_0 = y$,

Both the forward and the backward equations can be considered with different initial conditions. In particular, given a smooth function f, if we define

$$u(y,t) = \mathbb{E}_y f(X_t),$$

then $u(y,t) = \int_{\mathbb{R}} f(x)\rho(x,t|y)$ and hence it satisfies

$$\frac{\partial u}{\partial t} = b(y)\frac{\partial u}{\partial y} + \tfrac{1}{2}a(y)\frac{\partial^2 u}{\partial y^2},$$

with the initial condition $u(y,0) = f(y)$. In this sense, the SDE for X_t is the characteristic equation that is associated with this parabolic PDE, much in the same way as the ODE $\dot{X}_t = b(X_t)$ is the characteristic equation associated with the first order PDE $\partial u/\partial t = b(y)\partial u/\partial y$. This can be generalized in many ways. For instance, the solution of

$$\frac{\partial v}{\partial t} = c(y)v(y) + b(y)\frac{\partial v}{\partial y} + \tfrac{1}{2}a(y)\frac{\partial^2 v}{\partial y^2}.$$

with the initial condition $v(y,0) = f(y)$, can be expressed as

$$v(y,t) = \mathbb{E}_y f(X_t)e^{\int_0^t c(X_s)ds}.$$

This is the celebrated *Feynman-Kac formula* in the context of SDEs.

Let us consider an example. The forward differential equation associated with the Ornstein-Uhlenbeck process introduced in the last section is

$$\frac{\partial \rho}{\partial t} = \gamma\frac{\partial}{\partial x}(x\rho) + \frac{\sigma^2}{2}\frac{\partial^2 \rho}{\partial x^2}$$

The solution of this equation is

$$\rho(x,t|y) = \frac{1}{\sqrt{\pi\sigma^2(1-e^{-2\gamma t})/\gamma}} \exp\Big(-\frac{\gamma(x-ye^{-\gamma t})^2}{\sigma^2(1-e^{-2\gamma t})}\Big).$$

This shows that the Ornstein-Uhlenbeck process is a Gaussian process with mean $ye^{-\gamma t}$ and variance $\sigma^2(1-e^{-2\gamma t})/2\gamma$. It also confirms that this process tends to $N(0,\sigma^2/2\gamma)$ as $t \to \infty$ since

$$\rho(x) = \lim_{t\to\infty} \rho(x,t|y) = \frac{e^{-\gamma x^2/\sigma^2}}{\sqrt{\pi\sigma^2/\gamma}}.$$

Generally, the limit of $\rho(x,t|y)$ as $t \to \infty$, when it exists, gives the equilibrium density ρ of the process. It satisfies

$$0 = -\frac{\partial}{\partial x}(b(x)\rho) + \frac{1}{2}\frac{\partial^2}{\partial x^2}(a(x)\rho).$$

Forward and backward Kolmogorov equations can also be derived for multi-dimensional processes. They read respectively

$$\frac{\partial\rho}{\partial t} = -\sum_{j=1}^{J} \frac{\partial}{\partial x_j}(b_j(x)\rho) + \frac{1}{2}\sum_{j,j'=1}^{J} \frac{\partial^2}{\partial x_i \partial x_j}(a_{jj'}(x)\rho)$$

and

$$\frac{\partial\rho}{\partial t} = \sum_{j=1}^{J} b_j(x)\frac{\partial\rho}{\partial x_j} + \frac{1}{2}\sum_{j,j'=1}^{J} a_{jj'}(x)\frac{\partial^2\rho}{\partial x_i \partial x_j},$$

where $a_{jj'}(x) = \sum_{k=1}^{K} \sigma_{jk}(x)\sigma_{j'k}(x)$.

Notes by Walter Pauls and Arghir Dani Zarnescu.