Advanced computational methods X071521-SDE
Lecture 1

1 Motivation for SDEs

We consider a dynamics system with noise.

$$\dot{X} = b(t, X) + \sigma(t, X)\eta(t)$$

where $\eta(t)$ is the noise. For the noise $\eta(t)$, we have some intuition:

- $\mathbb{E}\eta(t) = 0$
- $\eta(t_2)$ and $\eta(t_1)$ have the same distribution.
- $\mathbb{E}\eta(t)\eta(s) = \delta(t - s)$.

The third condition means that the the noise at different times are unrelated. We make the correlation to be Dirac delta so that the noise has nontrivial contribution for the system. The Fourier transform of the correlation function

$$\mathcal{F}(\delta(\cdot)) = 1.$$

This is independent of the frequency and we thus call $\eta(t)$ the white noise.

Now, consider the integral of the white noise:

$$W(t) = \int_0^t \eta(s)\, ds.$$

This has the following properties

- $\mathbb{E}W = 0$
- $W(t)$ has stationary increments
- The increments are independent.
- $\mathbb{E}(W(t)W(s)) = s \wedge t$.

Moreover, if we expect W to have continuous paths, then $W(t)$ must be the Brownian motion, and

$$W(t) - W(s) \sim N(t - s).$$

This means that

$$\mathbb{E}W(t)W(s) = min(s, t).$$

Hence, the white noise is the derivative of Brownian motion.

Remark 1. *In the theory of random fields, the white noise is generalized to a $L^2(\Omega)$-valued measure (i.e. a mapping from Borel sets to L^2 random variables). Then, $B_t = \eta([0,t]) \in L^2(\Omega)$.*

Remark 2. *There is another Levy's theorem about Brownian motion that may be useful sometimes: If a martingale M has continuous paths and the quadratic variation is $[M,M] = t$, then M is the standard Brownian motion.*

With the understanding above, the dynamic system with noise can be written as

$$dX = b(t,X)dt + \sigma(t,X)dW,$$

or in integral form, we have

$$X(t) = X_0 + \int_0^t b(s, X(s))\, ds + \int_0^t \sigma(s, X(s))dW(s).$$

Such kind of equations are call stochastic differential equation (SDE) driven by Brownian motions. Of course, we have SDEs driven by other processes, which we will not touch in this course.

If σ is a constant, the integral of the last term gives $\sigma W(t)$, which is easy to understand. However, in general $\sigma(t,X)$ is random. We must understand how σ and dW are multiplied together. This will be answered by Itô integrals.

2 Itô integrals

The rigorous theory will be established on probability spaces using measure theory. For the measure space $(\Omega, \mathcal{F}, P)$. Here, $\mathcal{F}$ is the set (sigma algebra) of some subsets of Ω. These sets are called events. P is the probability measure. Each $\omega \in \Omega$ is called a sample point and you can understand that it corresponds to one realization of the Brownian motion.

2.1 Some basic definitions

A filtration $\{\mathcal{F}_t : t \in \mathbb{R}_+\}$ is a collection of σ-algebras on the probability space such that

$$\mathcal{F}_s \subset \mathcal{F}_t \subset \mathcal{F} \quad 0 \le s < t < \infty.$$

This models the increasing information of a system: $\mathcal{F}_t$ represents the current information we know.

A stochastic process $\{X_t\}$ is a family of random variables defined on $(\Omega, \mathcal{F}, P)$ indexed by t. Hence, $X = X(t,\omega)$. Alternatively, we can also understand it as a $[0,\infty) \to S$ function valued random variable $X(\cdot,\omega)$ on the probability space. Every such function is called a path.

We say X is adapted to the filtration $\{\mathcal{F}_t\}$ if $X(t) \in \mathcal{F}_t$. (This means

$$\{\omega : X(t,\omega) \in B\} \in \mathcal{F}_t, \text{for all Borel set } B)$$

Often, we choose the minimal filtration for the process

$$\mathcal{F}_t = \sigma\{X_s : 0 \le s \le t\}$$

This represents all the information brought by X up to time t.

For more rigorous and complete discussion of filtrations, stopping times, read the book draft written by Timo, available on his website.

Definition 1. *Given a process Y, the quadratic variation $[Y]$ is a stochastic process such that $t \mapsto [Y](t,\omega)$ is nondecreasing for all ω and*

$$[Y](t,\omega) = \lim_{mesh\to 0} \sum_i^n (Y_{t_{i+1}} - Y_{t_i})^2, \textit{ in probability}$$

Exercise: Prove that if $X(t)$ is a continuously differentiable, then $[X] = 0$. Prove that the quadratic variation is given by $[W] = t$.

Since Brownian motion has rough paths, the quadratic variation is nonzero. In fact,

$$[W] = t.$$

This formally means

$$(dW)^2 = dt.$$

This explains why in usual ODEs, we only have $df(X) = f'(X)dX$ but for diffusion processes the differential of $f(W)$ contains other terms.

The quadratic covariation is

$$[X,Y] = [\frac{1}{2}(X+Y)] - [\frac{1}{2}(X-Y)]$$

Intuitively, this is to define

$$\sum_i (X_{t_{i+1}} - X_{t_i})(Y_{t_{i+1}} - Y_{t_i})$$

The previous definition is used due to some path properities.

Hence, formally, we have

$$d[X,Y] = dXdY.$$

As we have seen, the quadratic variation is introduced because $(dW)^2$ has nontrivial contribution.

Remark 3. *In general, if we consider diffusion processes driven by Brownian motions, we need to consider quadratic variation and*

$$d(XY) = YdX + XdY + d[X, Y]$$

Here, YdX should be understood in Itô sense as we shall see soon.

Definition 2. *A real valued stochastic process $M = \{M_t : t \in \mathbb{R}_+\}$ is called a martingale with respect to the filtration $\{\mathcal{F}_t\}$ if M is adapted to the filtration and*

$$\mathbb{E}(M_t|\mathcal{F}_s) = M_s, \quad s < t.$$

Often, the martingales in $\mathcal{M}_2$ is good:

$$\|M\|_{\mathcal{M}_2} = \sum_{k=1}^{\infty} 2^{-k}(1 \wedge \|M_k\|_{L^2(\mathbb{P})}).$$

2.2 stochastic integrals

Suppose we want to define $\int_0^t WdW$. Consider the Riemann sum

$$S(\pi) = \sum_i W_{s_i}(W_{t_{i+1}} - W_{t_i})$$

where

$$s_i = (1-u)t_i + ut_{i+1}.$$

The computation starts with the following algebra identity

$$b(a-c) = \frac{1}{2}(a^2 - c^2) - \frac{1}{2}(a-c)^2 + (b-c)^2 + (a-b)(b-c)$$

Taking the sum, the first term is simply $\frac{1}{2}W_t^2$. The second term is the quadratic variation. The third term, by similar computation of quadratic variation, we can show that it converges in L^2 to ut:

$$\mathbb{E}(\sum_i (W_{s_i} - W_{t_i})^2) = \sum_i u(t_{i+1} - t_i) = ut,$$

$$Var(\sum_i (W_{s_i} - W_{t_i})^2) = \sum_i Var((W_{s_i} - W_{t_i})^2) = \sum_i 2(s_i - t_i)^2 \le 2tmesh(\pi) \to 0.$$

The third term has mean zero and variance converging to zero. Hence, the L^2 limit is given by

$$\frac{1}{2}W_t^2 - \frac{1}{2}t + ut.$$

We conclude the following

- The limit of the Riemann sum depends on the choice of sample point!
- If $u = 1/2$, we have the chain rule. However, $\frac{1}{2}W_t^2$ is not a martingale.
- If $u = 0$, we do not have chain rule but $\frac{1}{2}W_t^2 - \frac{t}{2}$ is a martingale.

If we choose the midpoint as the sample point, we get the Stratonovich integral. If we use the left point, the resulted integral is the Itô integral. Since the Itô integrals give martingales

Remark 4. *Show that if F has bounded total variation, then*

$$\lim_{mesh(\pi)\to 0} \sum_i F_{s_i}(F_{t_{i+1}} - F_{t_i})$$

is independent of the choice of u.

Remark 5. *For the stochastic integral with respect to cadlag semimartingales, the integrand should be predictable to make sense.*

Rigorous definition of Itô integral

The rigorous stochastic integral can be established for $X \in L^2([0,T]\times\Omega)$, i.e.

$$\|X\|^2_{L^2([0,T]\times\Omega)} := \mathbb{E}\int_{[0,T]} |X(t,\omega)|^2 dt < \infty.$$

- First of all, for the simple predictable process

  $$X_t(\omega) = \eta_0(\omega)1_{(0)}(t) + \sum_{i=1}^{n-1} \xi_i(\omega)1_{(t_i,t_{i+1}]}(t)$$

  where $0 = t_1 < t_2 < \ldots < t_n$. Predictable means that the state at t can be referred by the information in $s < t$. Also, ξ_i should be square integrable. The stochastic integral is defined by

  $$(X \cdot B)(t) = \int_0^t X dB = \sum_{i=1}^{n-1} \xi_i(\omega)(B_{t_{i+1}\wedge t}(\omega) - B_{t_i \wedge t}(\omega))$$

  Clearly, $X \cdot B$ is a martingale in $\mathcal{M}_2$ and we have the Itô isometry:

  $$\mathbb{E}[(X \cdot B)_t^2] = \mathbb{E}\int_0^t X_s^2\, ds, \quad \forall t \geq 0.$$

- The significant fact of the Itô isometry is that if $\{X_n\}$ is a sequence of simple predictable processes and $X_n \to X$ in $L^2([0,T]\times\Omega)$, then $X_n \cdot B$ is a Cauchy sequence in $\mathcal{M}_2$. Then, the limit is defined to be $(X\cdot B)(t) = \int_0^t X dB$.

  This is good enough because any $X \in L^2([0,T]\times\Omega)$ can be approximated by simple predictable processes.

  In the example above, $\sum_i B_{t_i} 1_{(t_i,t_{i+1}]}$ approximates B, and this corresponds to $u = 0$.

To summarize, for $X \in L^2([0,T]\times\Omega)$, we can define $X\cdot B = \int_0^t X dB$ which is a martingale and the Itô isometry holds

$$\mathbb{E}(\int_0^t X dB)^2 = \mathbb{E}\int_0^t X_s^2\, ds$$

There are extensions to processes that are not in L^2. Those who are interested can read the reference.

3 Itô formula

As we have seen, the Brownian motion has rough paths and $d[B,B] = dt$. Intuitively, this means $(dB)^2 = dt$. Hence, if we expand $f(B)$, the quadratic variation term will be nontrivial. Then, we have

$$f(B) = f(B_0) + \int_0^t f'(B_s)\, dB_s + \frac{1}{2}\int_0^t f''(B_s)\, d[B]_s.$$

For standard Brownian motion $W_0 = 0$. However, we can also consider non-standard Brownian motion

$$B(t) = B_0 + W(t).$$

That is why we have B_0.

This is the special case of the Itô's formula. The most general Itô's formula is valid for stochastic integrals with respect to semi-martingales.

Here, I only list the case for Brownian motions in $\mathbb{R}^d$.

Theorem 1. *Let $B = (B_1, B_2, \ldots, B_d)$ be a Brownian motion in $\mathbb{R}^d$ with random initial data $B(0)$. Let $f \in C^2(\mathbb{R}^d)$. Then, we have*

$$f(B(t)) = f(B(0)) + \int_0^t \sum_i \partial_i f(B_s) dB_s + \frac{1}{2}\int_0^t \Delta f(B_s)\, ds.$$

We will omit the proof. To memorize this, you can understand it as

$$dB_i dB_j = \delta_{ij} dt, \quad dB_i dt = 0, \quad (dB_i)^p = 0, p \geq 3.$$

4 Stochastic differential equations

The general stochastic differential equations are given by

$$dX = dH + F(t, X)dY$$

where Y is a general cadlag semimargingales. In this course, we only focus on the Itô equations

$$dX = b(t, X)dt + \sigma(t, X)dB_t, \quad X_0 = \xi.$$

This equation is defined by the following integral equation

$$X(t) = \xi + \int_0^t b(s, X_s)\, ds + \int_0^t \sigma(s, X_s) dB_s.$$

The integral is in the Itô sense.

b is $\mathbb{R}^d$ valued and is called the drift vector. σ is called the dispersion matrix which has size $d \times m$. The matrix $\sigma\sigma^T$ is called the diffusion matrix. The solution of Itô equations will be called the diffusion processes.

Before we go to the rigorous theory, let us look at two examples.

Example 1: the Ornstein-Uhlenbeck process

$$dX = -\alpha X dt + \sigma dW_t.$$

Assume the initial data X_0 is independent of the Brownian motion.

Mimicking the technique for ODE, we want to try integrating factor. However, the processes we have all nontrivial quadratic variation.

$$d(ZX) = ZdX + Xdz + d[Z, X]$$

Let us try

$$Z = \exp(\alpha t)$$

Then, $[Z, X] = 0$ because $dBdt = 0$. Then,

$$d(ZX) = -\alpha ZX dt + \sigma ZdB + \alpha ZX dt = \sigma ZdB.$$

Hence, we in fact have the usual formula as in ODE.

The OU process is then solved to be

$$X_t = X_0 e^{-\alpha t} + \sigma e^{-\alpha t} \int_0^t e^{\alpha s} dW_s.$$

Of course, this is formal guess, you may need to verify that it satisfies the integral equation by Itô formula, which is left for exercise.

Exercise: Compute the mean and the variance of the 1D OU process. Since the dispersion matrix does not depend on the process X, X is a Gaussian process, write out the density of X_t.

Example 2. Geometric Brownian motion

$$dX = \mu X dt + \sigma X dB, \quad X(0) = x_0.$$

For the integrating factor, one may guess to use

$$Z = \exp(-\mu t - \sigma B_t)$$

and get

$$X = X_0 \exp(\mu t + \sigma B_t)$$

This turns out to be wrong. In fact,

$$d(XZ) = XdZ + ZdX + d[Z, X] = \frac{1}{2} XZ\sigma^2 dt + d[Z, X]$$

The quadratic variation part is nonzero:

$$d[X, Y] = -\sigma^2 XZdt.$$

Hence,

$$d(XY) = -\frac{1}{2} XZ\sigma^2 dt.$$

What is the correct integrating factor? Motivated by the above computation, we try into the factor

$$Z = \exp(-\mu t - \sigma B_t + r\sigma^2 t)$$

Then,

$$d(XZ) = XdZ + ZdX + d[X, Z] = (\frac{1}{2} + r)\sigma^2 XZdt - \sigma^2 XZdt.$$

Clearly, we need $r = \frac{1}{2}$.

Hence, the geometric Brownian motion should be solved as

$$X_t = X_0 \exp\left((\mu - \frac{1}{2})\sigma^2 t + \sigma B_t \right).$$

To verify this is a solution, we need to check all the assumptions in the derivation above are valid. Alternatively, one can check directly by inserting this into the integral equation.

Exercise: Use Itô's formula to find an ODE for $u(t) = \mathbb{E}X^2$ for the geometric Brownian motion. Then, find $u(t)$.