Recitation 4: Stochastic Integrals, SDEs, and their simulation.

Recap: A stochastic differential equation

$$\dot{x} = f(x) + \sigma(x)\,\eta(t)$$

has solutions that locally resemble a generalized Brownian motion:

$$x(t) \simeq x(t_0) + f(x(t_0))(t-t_0) + \sigma(x(t_0))\,W(t-t_0), \quad \text{as } t \to t_0$$

Stochastic Integrals

A stochastic differential equation is a shorthand for the corresponding stochastic integral:

$$x(t) = \int_0^t ds\, f(x(s)) + \int_0^t ds\, \sigma(x(s))\,\eta(s)$$

In math, integrals of the form $\int_{t_0}^t ds\,\eta(s)\,h(s)$ are typically denoted

$$\int_{t_0}^t ds\,\eta(s)\,h(s) = \int_{t_0}^t dW(s)\,h(s) \quad (I)$$

i.e. we replace $ds\,\eta(s) \longrightarrow dW(s)$.

The corresponding notation for an SDE is

$$dx = f(x,t)\,dt + \sigma(x,t)\,dW$$

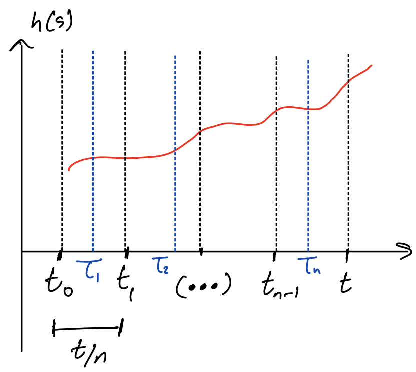
We define (I) as a stochastic version of the Riemann–Stieltjes integral:

Partition the interval $[t_0, t]$ into n subintervals,

$$t_0 \le t_1 \le \cdots \le t_{n-1} \le t$$

Pick times τ_i within these subintervals:

$$t_{i-1} \le \tau_i \le t_i$$



The stochastic integral is defined as the limit

$$\int_{t_0}^t h(s)\,dW(s) \equiv \lim_{n\to\infty} S_n,$$

where

$$S_n \equiv \sum_{i=1}^n h(\tau_i)\left[W(t_i) - W(t_{i-1})\right]$$

and W is a standard Wiener process.

Note that the value of this integral depends on the choice of τ_i.
For example, consider $h(s) = W(s)$. Then:

$$\langle S_n \rangle = \sum_{i=1}^{n} \langle W(\tau_i)[W(t_i) - W(t_{i-1})] \rangle$$
$$= \sum_{i=1}^{n} [\min(\tau_i, t_i) - \min(\tau_i, t_{i-1})]$$
$$= \sum_{i=1}^{n} (\tau_i - t_{i-1})$$

If we use an "α-discretization" as mentioned in lecture:

$$\tau_i \equiv \alpha t_i + (1-\alpha) t_{i-1}, \qquad 0 \le \alpha \le 1,$$

Then

$$\langle S_n \rangle = \sum_{i=1}^{n} \alpha(t_i - t_{i-1}) = \alpha(t - t_0)$$

This ranges anywhere between 0 and $t - t_0$.

In the Itō perscription, we make the choice

$$\alpha = 0 \quad \leftarrow \text{we primarily use this.}$$

In the Stratonovich perscription, we use

$$\alpha = 1/2.$$

For arbitrary $h(s)$, there is no general correspondence between the two integrals, but in the important special case where $h(s) = h(x(s))$ and $x(s)$ is the solution to an SDE, there is in fact a general formula relating the two perscriptions. We will come back to this in a future recitation.

1) Example: exact calulation of $\int_0^t W(s)\, dW(s)$ (in the Ito sense)

We work in the Ito perscription and let $W_i \equiv W(t_i)$, $\Delta W_i \equiv W_i - W_{i-1}$

$$S_n = \sum_{i=1}^{n} W_{i-1} \Delta W_i$$
$$= \frac{1}{2} \sum_{i=1}^{n} [(W_{i-1} + \Delta W_i)^2 - W_{i-1}^2 - \Delta W_i^2]$$
$$= \frac{1}{2} \sum_{i=1}^{n} [W_i^2 - W_{i-1}^2] - \frac{1}{2} \sum_{i=1}^{n} \Delta W_i^2$$
$$= \frac{1}{2} W(t)^2 - \frac{1}{2} \sum_{i=1}^{n} \Delta W_i^2$$

The sum $\sum_{i=1}^{n} \Delta W_i^2$ is the so called "Quadratic variation" of the Wiener process.

Since the terms are i.i.d, we may invoke the Law of Large numbers as $n \to \infty$:

$$\sum_{i=1}^{n} \Delta W_i^2 \longrightarrow \sum_{i=1}^{n} \langle \Delta W_i^2 \rangle = \sum_{i=1}^{n} \frac{t}{n} = t.$$

Thus:

$$\boxed{\int_0^t W(s)\, dW(s) = \frac{1}{2}\left(W_t^2 - t\right)}$$

2) Expectation values of Ito integrals

Let us first define the notion of an "adapted" or "nonanticipating" function.

Def: A stochastic process M is adapted to W if, for every t, $M(t)$ is some deterministic function of the portion $\{W(s) : 0 \le s \le t\}$ of the path of W up to time t.

i.e. $M(t)$ is determined by the history of W and is independent of its future.

⇒ For example, W is itself adapted to W.

⇒ The solution to an SDE $x(t) = x_0 + \int_0^t ds\, f(x(s)) + \int_0^t dW(s)\, \sigma(x(s))$ is also adapted to W.

Ito integrals of adapted processes have an important property:

Thm: if M is adapted to W, then

$$\left\langle \int_{t_0}^{t} M(s)\, dW(s) \right\rangle = 0.$$

Pf: $\left\langle \int_{t_0}^{t} M(s)\, dW(s) \right\rangle = \lim_{n\to\infty} \langle S_n \rangle$

(notice that this does not apply to Stratonovich integrals)

$$\langle S_n \rangle = \sum_{i=1}^{n} \langle M(t_{i-1})\, \Delta W_i \rangle$$

If M is adapted to W, then it can be written as a deterministic function of increments of W within the interval $[0, t_{i-1}]$. By the independence of nonoverlapping increments, we conclude $\langle M(t_{i-1}) \Delta W_i \rangle = \langle M(t_{i-1}) \rangle \langle \Delta W_i \rangle = 0.$ ∎

This means that if $x(t)$ is an Ito process, all integrals of the form

$$\boxed{\int_0^t g(x(s))\, dW(s)} \quad \text{for any function } g$$

have vanishing expectation. This is consistent with the notation in lecture: setting $dW(s) = \eta(s) ds$, we say that $\langle g(x(s))\, \eta(s) \rangle = 0$.

You can check the theorem for $\int W(s) dW(s)$ calculated explicitly above.

Simulating an Ito process (SDE)

1) The Euler-Maruyama algorithm.

From now on we work in the Ito convention. We wish to numerically sample the solutions to an SDE:

$$\dot{x} = f(x,t) + \sigma(x,t)\,\eta(t)$$

That is, we wish to sample the stochastic integral

$$x(t) = \int_0^t ds\, f(x(s),s) + \int_0^t dW(s)\, \sigma(x(s),s)$$

The simplest way to do this is to directly apply the definition of the Ito integral by discretizing time $t_k \equiv k\Delta t$:

$$x(t_n) \equiv x(0) + \sum_{k=0}^{n-1} f(x(t_k),t_k)\,\Delta t + \sum_{k=0}^{n-1} \sigma(x(t_k),t_k)\left[W(t_{k+1}) - W(t_k)\right]$$

The RHS depends only on the past values $\{x(t_0),\dots,x(t_{n-1})\}$, and on the quantity

$$W(t_k+\Delta t) - W(t_k) \sim \mathcal{N}(0,\sqrt{\Delta t}).$$

$x(t_n)$ can thus be determined recursively:

$$x(t_{n+1}) = x(t_n) + f(x(t_n),t_n)\,\Delta t + \sigma(x(t_n),t_n)\sqrt{\Delta t}\; Z_n$$

where the $\{Z_n\}$ are i.i.d $\mathcal{N}(0,1)$.

This defines the so-called Euler-Maruyama (EM) algorithm.

(1) Set $x(0) = x_0$, $t=0$.

(2) Draw $Z \sim \mathcal{N}(0,1)$

(3) Set $x(t+\Delta t) = f(x(t),t)\,\Delta t + \sigma(x(t),t)\sqrt{\Delta t}\; Z$

(4) Set $t = t+\Delta t$. If $t < t_{max}$, return to (2).

This algorithm becomes more accurate as $\Delta t \to 0$. How quickly does it converge? For multiplicative noise, we will show shortly that the error scales as $\sqrt{\Delta t}$.

This is much worse than the standard Euler scheme for an ODE, for which the error is $\mathcal{O}(\Delta t)$.

You may know that for ODEs, there exist higher order methods methods that converge faster than $O(\Delta t)$. There are similar methods for SDEs, but it is a much trickier business than for ODEs, in part because of the variety of definitions one can use to quantify the convergence of SDE integrators.

A detailed account of such techniques can be found in the book:

Kloeden, P. E.; Platen, E, "Numerical Solution of Stochastic Differential Equations", Springer-Verlag (1992).

We discuss below a procedure for generating integration schemes that improve on EM.

2) The Stochastic Ito-Taylor expansion

a) Warm up: Deterministic case

Let us first consider deterministic ODEs of the form:

$$\dot{x} = f(x)$$

We write this in integral form

$$x(t) = x(0) + \int_0^t f(x(s))\, ds \qquad (2)$$

Now note that for any function $g(x)$, we have $\dot{g}(x) = g'(x)\,\dot{x} = f(x)\, g'(x)$, so that

$$g(x(t)) = g(x(0)) + \int_0^t ds\, f(x(s))\, g'(x(s))$$
$$\equiv g(x(0)) + \int_0^t ds\, \mathcal{L} g\,(x(s)), \qquad \text{where } \mathcal{L} \equiv f(x)\,\partial_x$$

Let us now use this to rewrite the integral in (2) by setting $g = f$:

$$x(t) = x(0) + f(x(0)) \int_0^t ds + \int_0^t ds \int_0^s du\, \mathcal{L} f(x(u))$$
$$= x(0) + f(x(0))\, t + R_2,$$

where $R_2 \equiv \int_0^t ds \int_0^s du\, \mathcal{L} f(x(u)) \in O(t^2)$. For small t, we may thus neglect R_2.

This gives the Euler scheme $x(t) = x(0) + f(x(0))\,t + R_2$.

To do better, we continue the expansion, setting $g = \mathcal{L} f$:

$$x(t) = x(0) + f(x(0))\, t + \mathcal{L} f(x(0)) \int_0^t ds \int_0^s du + \underline{\int_0^t ds \int_0^s du \int_0^u dv\, \mathcal{L}^2 g(x(v))}$$
$$\equiv x(0) + f(x(0))\, t + \tfrac{1}{2} \mathcal{L} f(x(0))\, t^2 + R_3.$$

We may in general work up to any order:

$$x(t) = x(0) + \sum_{m=1}^{r} \frac{t^m}{m!} \mathcal{L}^{m-1} f(x(0)) + R_{r+1}$$

where $R_{r+1} \in \mathcal{O}(t^{r+1})$ is

$$R_{r+1} = \int_0^t ds_1 \int_0^{s_1} ds_2 \cdots \int_0^{s_r} ds_{r+1} \, \mathcal{L}^r f(x(s_{r+1}))$$

This is effectively a Taylor expansion in integral form.

b) Stochastic case and the Milstein scheme

In moving to SDEs, the main difference is that we must use Ito's formula in place of the chain rule. For an SDE

$$\dot{x} = f(x) + \sigma(x) \, \eta(t),$$

we have

$$\begin{aligned} \dot{g}(x) &= g'(x) \, \dot{x} + \tfrac{1}{2} \sigma(x) \, g''(x) \\ &= g'(x) f(x) + \tfrac{1}{2} \sigma(x) g''(x) + g'(x) \sigma(x) \, \eta(t) \end{aligned}$$

Integrating:

$$g(x(t)) = g(x(0)) + \int_0^t ds \, \mathcal{L}_0 \, g(x(s)) + \int_0^t \mathcal{L}_1 \, g(x(s)) \, dW(s).$$

where

$$\mathcal{L}_0 \equiv f(x) \, \partial_x + \tfrac{1}{2} \sigma(x) \, \partial_x'', \qquad \mathcal{L}_1 \equiv \sigma(x) \, \partial_x$$

we now apply this to

$$x(t) = x(0) + \int_0^t ds \, f(x(s)) + \int_0^t \sigma(x(s)) \, dW(s).$$

Setting $g = f$ and $g = \sigma$ in the above:

$$x(t) = x(0) + f(x(0)) \underbrace{\int_0^t ds}_{t} + \sigma(x(s)) \underbrace{\int_0^t dW(s)}_{W(t)} + R_1$$

where

$$R_1 \equiv \underbrace{\int_0^t ds \int_0^s du \, \mathcal{L}_0 f(x(u))}_{\mathcal{O}(t^2)} + \overbrace{\int_0^t ds \int_0^s dW(u) \, \mathcal{L}_1 g(x(s))}^{\mathcal{O}(t^{3/2})}$$

$$+ \overbrace{\int_0^t dW(s) \int_0^s du \, \mathcal{L}_0 \sigma(x(u))}^{\mathcal{O}(t^{3/2})} + \overbrace{\int_0^t dW(s) \int_0^s dW(u) \, \mathcal{L}_1 \sigma(x)}^{\mathcal{O}(t)}$$

If we set $R_1 = 0$, we get the EM scheme. This is accurate only to $\mathcal{O}(\sqrt{t})$ because we have neglected a term of order t in R_1. To get an algorithm that's accurate to $\mathcal{O}(t)$, we must further expand this term:

$$\int_0^t dW(s)\int_0^s dW(u)\ \mathcal{L}_1\sigma(x) = \mathcal{L}_1\sigma(x(0))\int_0^t dW(s)\int_0^s dW(u) + o(t)$$

$$= \mathcal{L}_1\sigma(x_0)\int_0^t W(s)\,dW(s) \longleftarrow \text{we calculated this above}$$

$$= \frac{\sigma(x_0)\sigma'(x_0)}{2}(W_t^2 - t)$$

To conclude, we have

$$x(t) = x_0 + f(x_0)t + \sigma(x_0)W(t) + \tfrac{1}{2}\sigma(x_0)\sigma'(x_0)(W_t^2 - t) + R_2$$

where $R_2 \in O(t^{3/2})$. Setting $R_2 = 0$ gives the Milstein scheme, accurate to $O(\Delta t)$:

Def Milstein Scheme: to integrate an SDE $\dot{x} = f(x) + \sigma(x)\eta(t)$, discretize time $t_n = n\Delta t$, and compute $x(t_n) \equiv x_n$ recursively as

$$x_{n+1} = x_n + f(x_n)\Delta t + \sigma(x_n)\sqrt{\Delta t}\,Z_n + \tfrac{1}{2}\sigma(x_n)\sigma'(x_n)(\Delta t\, Z_n - t)$$

where $\{Z_n\}$ are i.i.d. $\mathcal{N}(0,1)$.

Note that for additive noise $\sigma(x) = \sigma$, the Milstein scheme reduces to the EM scheme, implying that EM is also accurate to $O(\Delta t)$ in the additive case.