Stochastic Calculus

Week 4

Topics:

- Ordinary (Newtonian) Calculus
- The Riemann-Stieltjes Integral
- The Itô Integral – an Example
- Construction of the Itô Integral
- Itô processes and Itô's lemma

Ordinary (Newtonian) Calculus

Let $F(t)$ be a differentiable function, such that its derivative

$$\frac{dF(t)}{dt} = \lim_{h\to 0} \frac{F(t+h) - F(t)}{h}$$

exists. Then the following three statements are equivalent:

1. $\frac{dF(t)}{dt} = f(t)$ (definition of derivative)
2. $dF(t) = f(t)dt$ (differential equation)
3. $F(t) = F(0) + \int_0^t f(s)ds$ (integral equation)

This gives an easy way to interpret integrals like

$$\int_0^t g(s)dF(s) = \int_0^t g(s)f(s)ds.$$

In stochastic calculus, we want to consider similar expressions with $F(t)$ replaced by a Brownian motion process B_t. However, since Brownian sample paths are not differentiable, the above intuition will not work: there is no process $B_t' = dB_t/dt$, and hence we cannot express $\int_0^t g(s)dB_s$ as $\int_0^t g(s)B_s'ds$.

The Riemann-Stieltjes Integral

Let f and g both be functions of $t \in [0, T]$, and consider a partition $0 = t_0 < t_1 < \ldots < t_n = T$, and a corresponding set of intermediate time points values t_i^* (called y_i by Mikosch) such that $t_{i-1} \le t_i^* \le t_i$. We assume that $\max |t_i - t_{i-1}| \to 0$ as $n \to \infty$. Then the *Riemann-Stieltjes integral* of f with respect to g is

$$\int_0^T f(t)dg(t) = \lim_{n\to\infty} \sum_{i=1}^n f(t_i^*)[g(t_i) - g(t_{i-1})],$$

if the limit exists. When $g(t) = t$, we call this a *Riemann integral*.

A useful result is *integration by parts*, obtained as follows. Let $t_1^* = t_0 = 0$ and $t_{n+1}^* = t_n = T$. Write the sum as

$$\begin{aligned}\sum_{i=1}^n f(t_i^*)g(t_i) - \sum_{i=1}^n f(t_i^*)g(t_{i-1}) \\ = f(t_{n+1}^*)g(t_n) - f(t_1^*)g(t_0) \\ - \sum_{i=1}^n g(t_i)[f(t_{i+1}^*) - f(t_i^*)],\end{aligned}$$

which converges to $f(T)g(T) - f(0)g(0) - \int_0^T g(t)df(t)$.

The Riemann-Stieltjes integral exists if either f or g (or both) are differentiable / of bounded variation.

Example: Let $f(t) = e^{ct}$ and $g(t) = B_t$. Then f is differentiable, but g is not. In that case we can simply define the integral $\int_0^T e^{ct}dB_t$, via integration by parts as

$$\begin{aligned}\int_0^T e^{ct}dB_t &= \int_0^T f(t)dB_t \\ &= f(T)B_T - f(0)B_0 - \int_0^T B_t df(t) \\ &= e^{cT}B_T - \int_0^T B_t de^{ct} \\ &= e^{cT}B_T - \int_0^T cB_t e^{ct}dt.\end{aligned}$$

The final integral is an ordinary Riemann integral of a function of a Brownian motion, which is defined “pathwise”.

Note: Riemann integrals of functions of Brownian motions are stochastic, but they do not fall in the class of *stochastic integrals*; the latter name is reserved for so-called Itô integrals of the form $\int_0^T f(B_t, t)dB_t$.

The Itô Integral – an Example

Suppose we wish to define $\int_0^t B_s dB_s$, and we try to do this via the Riemann-Stieltjes construction. For simplicity take $t_i = it/n$. Let us take $t_i^* = t_{i-1} = (i-1)t/n$ as the intermediate points for the moment. Then we will have to find the limit of

$$\sum_{i=1}^n B_{t_i^*}[B_{t_i} - B_{t_{i-1}}] = \sum_{i=1}^n B_{t_{i-1}}[B_{t_i} - B_{t_{i-1}}].$$

Define $\Delta_i B = B_{t_i} - B_{t_{i-1}}$, and note that

$$\begin{aligned} B_t^2 &= \left(\sum_{i=1}^n \Delta_i B\right)^2 \\ &= \sum_{i=1}^n (\Delta_i B)^2 + 2\sum_{i=2}^n \sum_{j=1}^{i-1} (\Delta_i B)(\Delta_j B) \\ &= \sum_{i=1}^n (\Delta_i B)^2 + 2\sum_{i=2}^n B_{t_{i-1}} \Delta_i B, \end{aligned}$$

so that we are now interested in finding the limit of

$$\sum_{i=1}^n B_{t_{i-1}} \Delta_i B = \frac{1}{2}\left(B_t^2 - \sum_{i=1}^n (\Delta_i B)^2\right).$$

The random variables $\Delta_i B$ are i.i.d. $N(0, t/n)$, so that $Z_i := \sqrt{n/t}\,\Delta_i B \sim$ i.i.d. $N(0,1)$. One can show that the following convergence

$$\sum_{i=1}^n (\Delta_i B)^2 = \frac{t}{n}\sum_{i=1}^n Z_i^2 \to t\mathbb{E}(Z_i^2) = t$$

holds *in mean square*, meaning that

$$\mathbb{E}\left[\left(\sum_{i=1}^n (\Delta_i B)^2 - t\right)^2\right] \to 0.$$

(See Mikosch, p.98, for a proof).

This means that we may define the $\int_0^t B_s dB_s$ as the following limit in mean square:

$$\begin{aligned} \int_0^t B_s dB_s &= \lim_{n\to\infty} \sum_{i=1}^n B_{t_{i-1}}[B_{t_i} - B_{t_{i-1}}] \\ &= \frac{1}{2}\left(B_t^2 - t\right). \end{aligned}$$

Note that if we would have taken $t_i^* = t_i$ instead of t_{i-1}, then a different result would obtain:

$$\begin{aligned}\sum_{i=1}^n B_{t_i}[B_{t_i} - B_{t_{i-1}}] &= \sum_{i=1}^n B_{t_{i-1}}[B_{t_i} - B_{t_{i-1}}] + \sum_{i=1}^n (\Delta_i B)^2 \\ &\to \frac{1}{2}\left(B_t^2 - t\right) + t \\ &= \frac{1}{2}\left(B_t^2 + t\right).\end{aligned}$$

This shows that the integral $\int_0^t B_s dB_s$ does not exist in Riemann-Stieltjes sense: it matters how we choose $t_i^* \in [t_{i-1}, t_i]$. The limit is of a "stochastic" type, and does not exist "pathwise".

Note: An important difference between $\frac{1}{2}(B_t^2 - t)$ and $\frac{1}{2}(B_t^2 + t)$ is that the former is a martingale, whereas the latter is not. This martingale property comes from taking $t_i^* = t_{i-1}$, i.e., at the beginning of the interval $[t_{i-1}, t_i]$ over which we calculate the increments $\Delta_i B$. This implies

$$\begin{aligned}&\mathbb{E}(B_{t_i^*}[B_{t_i} - B_{t_{i-1}}]|\mathcal{F}_{t_{i-1}}) \\ &\quad = B_{t_i^*}\mathbb{E}([B_{t_i} - B_{t_{i-1}}]|\mathcal{F}_{t_{i-1}}) = 0.\end{aligned}$$

Construction of the Itô Integral

We start by defining a *simple process* C_t on $[0, T]$. Consider a partition $0 = t_0 < \ldots < t_n = T$ as usual, and let $\{Z_i\}$ be a sequence of random variables such that Z_i is a function of $(B_s, s \leq t_{i-1})$, i.e., adapted to $\mathcal{F}_{t_{i-1}} = \sigma(B_s, s \leq t_{i-1})$. Also assume $\mathbb{E}(Z_i^2) < \infty$. Then define

$$C_t = \sum_{i=1}^n I_{[t_{i-1}, t_i)}(t) Z_i + I_{\{t_n\}}(t) Z_n,$$

or in other words

$$C_t = \begin{cases} Z_i & \text{if} \quad t_{i-1} \leq t < t_i, \\ Z_n & \text{if} \quad t = T. \end{cases}$$

Then we *define*

$$\begin{aligned}\int_0^T C_s dB_s &:= \sum_{i=1}^n C_{t_{i-1}}(B_{t_i} - B_{t_{i-1}}) \\ &= \sum_{i=1}^n Z_i \Delta_i B = Z \cdot \Delta B.\end{aligned}$$

Note that this is a random variable. We can extend this definition to a stochastic process by letting the endpoint t of the integral vary from 0 to T:

$$
\begin{aligned}
X_t &:= \int_0^t C_s dB_s \\
&= \sum_{i=1}^{k-1} Z_i \Delta_i B + Z_k(B_t - B_{t_{k-1}}), \qquad t_{k-1} < t \le t_k.
\end{aligned}
$$

Two important properties of this process:

1. $\{X_t, \mathcal{F}_t\}$ is a martingale. Suppose that $t_{l-1} \le s < t_l$ and $t_{k-1} \le t < t_k$. Then

$$
\begin{aligned}
X_t - X_s &= Z_l(B_{t_l} - B_s) \\
&\quad + \sum_{i=l+1}^{k-1} Z_i \Delta_i B \\
&\quad + Z_k(B_t - B_{t_{k-1}}).
\end{aligned} \tag{1}
$$

   Each of the right-hand side terms has conditional expectation zero given $\mathcal{F}_s$. This may be shown using the "tower property". For example,

$$
\begin{aligned}
\mathbb{E}(Z_k(B_t - B_{t_{k-1}})|\mathcal{F}_s) &= \mathbb{E}[\mathbb{E}(Z_k(B_t - B_{t_{k-1}})|\mathcal{F}_{t_{k-1}})|\mathcal{F}_s] \\
&= \mathbb{E}[Z_k \mathbb{E}((B_t - B_{t_{k-1}})|\mathcal{F}_{t_{k-1}})|\mathcal{F}_s] \\
&= 0.
\end{aligned}
$$

   The other terms follow analogously.
   Thus $\mathbb{E}(X_t - X_s|\mathcal{F}_s) = 0$, or $\mathbb{E}(X_t|\mathcal{F}_s) = X_s$.

2. $\mathbb{E}[(X_t - X_s)^2|\mathcal{F}_s] = \mathbb{E}\left[\int_s^t C_u^2 du|\mathcal{F}_s\right]$. The proof of this property follows from taking the conditional expectation of the square of (1), and noting that all cross-products have zero conditional expectation; for $l \le i < j \le k$:

$$
\begin{aligned}
\mathbb{E}(Z_i \Delta_i B Z_j \Delta_j B|\mathcal{F}_s) &= \mathbb{E}[Z_i \Delta_i B Z_j \mathbb{E}(\Delta_j B|\mathcal{F}_{t_{j-1}})|\mathcal{F}_s] \\
&= 0.
\end{aligned}
$$

   For the squares, we find

$$
\begin{aligned}
\mathbb{E}(Z_i^2 \Delta_i B^2|\mathcal{F}_s) &= \mathbb{E}[Z_i^2 \mathbb{E}(\Delta_i B^2|\mathcal{F}_{t_{i-1}})|\mathcal{F}_s] \\
&= \mathbb{E}[Z_i^2 (t_i - t_{i-1})|\mathcal{F}_s],
\end{aligned}
$$

   so that (taking $s = t_{l-1}$ and $t = t_{k-1}$ for simplicity)

$$
\begin{aligned}
\mathbb{E}[(X_t - X_s)^2|\mathcal{F}_s] &= \mathbb{E}\left[\sum_{i=l}^{k-1} Z_i^2 (t_i - t_{i-1})|\mathcal{F}_s\right] \\
&= \mathbb{E}\left[\int_s^t C_u^2 du|\mathcal{F}_s\right].
\end{aligned}
$$

   Setting $s = 0$, this yields $\mathrm{var}(X_t) = \mathbb{E}(\int_0^t C_s^2 ds)$. This property is referred to as the *Itô isometry*.

3. If C_t is non-stochastic, then $X_t \sim N\left(0, \int_0^t C_s^2 ds\right)$. This follows since X_t is essentially a weighted sum of Brownian increments, which are independent and normally distributed.

4. The *quadratic variation* of a martingale X_t, denoted $\langle X \rangle_t$, is the probability limit, for ever finer partitions, of
$$\sum_{i=1}^{k-1} (X_{t_i} - X_{t_{i-1}})^2 + (X_t - X_{t_{k-1}})^2,$$
where again $t_{k-1} \leq t < t_k$. We sometimes denote it as $\langle X \rangle_t = \int_0^t (dX_s)^2$, or $d\langle X \rangle_t = (dX_t)^2$. For Brownian motion, $\langle B \rangle_t = t$; for an integral $X_t = \int_0^t C_s dB_s$ of a simple process, it is $\langle X \rangle_t = \int_0^t C_s^2 ds$.

So far we have only dealt with integrals of simple processes. Suppose now that C_t is a more general process on $[0, T]$, adapted to $\mathcal{F}_t = \sigma(B_s, s \leq t)$ and satisfying $E(\int_0^T C_t^2 dt) < \infty$. The construction of a general Itô integral is as follows:

1. C_t may be seen as a limit, in mean square, of a sequence of adapted simple processes $C_t^n, n = 1, 2, \ldots$. That is:
$$\mathbb{E}\left[\int_0^T (C_t^n - C_t)^2 dt\right] \to 0.$$
This result will not be proved here. In practice we will take the approximating simple processes as
$$C_t^n = \sum_{i=1}^n I_{[t_{i-1}, t_i)}(t) C_{t_{i-1}} + I_{\{t_n\}}(t) C_{t_n}$$

2. The "distance" (in the mean square sense) of two stochastic integrals $X_t^n = \int_0^t C_s^n dB_s$ and $X_t^m = \int_0^t C_s^m dB_s$ is
$$\begin{aligned} & \mathbb{E}\left[\left(\int_0^T C_s^n dB_s - \int_0^T C_s^m dB_s\right)^2\right] \\ = \; & \mathbb{E}\left[\left(\int_0^T [C_s^n - C_s^m] dB_s\right)^2\right] \\ = \; & \mathbb{E}\left[\left(\int_0^T (C_s^n - C_s^m)^2 ds\right)^2\right], \end{aligned}$$

and this distance will go to zero as $m, n \to \infty$, because both C_t^n and C_t^m converge to C_t.

3. Because the distance between X_t^n and X_t^m converges to zero, there must be some limit point X_t where both X_t^n and X_t^m converge to. We'll call this limit point $X_t = \int_0^t C_s dB_s$. Thus, it is defined by the limit in mean square of
$$\int_0^T C_s^n dB_s = \sum_{i=1}^{n} C_{t_{i-1}}(B_{t_i} - B_{t_{i-1}}).$$
This definition also gives the integral at time t instead of T, by replacing C_s by $C_s I_{[0,t]}(s)$.

4. This general Itô integral inherits the most important properties of the integral of a simple process: X_t is a martingale, with quadratic variation process $\langle X \rangle_t = \int_0^t (dX_s)^2 = \int_0^t C_s^2 ds$. Thus its variance is also $\mathrm{var}(X_t) = E(\int_0^t C_s^2 ds)$, and if C_t is non-stochastic, then
$$X_t = \int_0^t C_s dB_s \sim N\left(0, \int_0^t C_s^2 ds\right).$$

Itô Processes and Itô's Lemma

We now have given meaning to the integral equation $X_t = \int_0^t \sigma_s dB_s$, where σ_s satisfies the same properties as C_s before. Now let μ_t be another adapted process (a function of $B_s, s \leq t$), and define
$$X_t = X_0 + \int_0^t \mu_s ds + \int_0^t \sigma_s dB_s.$$
Then X_t is called an Itô process. We call μ_t the *drift* of the process, and σ_t the *volatility*. Note that the first integral is an ordinary Riemann integral, whereas the second is a stochastic integral.

The integral equation may be re-expressed in differential form:
$$dX_t = \mu_t dt + \sigma_t dB_t.$$
This is just another way of writing the integral equation; the meaning of a differential term like $\sigma_t dB_t$ stems from the corresponding stochastic integral.

Interpretation of μ_t and σ_t: consider the increment of X_t over $[t, t+h]$:

$$X_{t+h} - X_t = \int_t^{t+h} \mu_s ds + \int_t^{t+h} \sigma_s dB_s.$$

Since the stochastic integral is a martingale, we have that

$$\begin{aligned} \mathbb{E}\left(X_{t+h} - X_t | \mathcal{F}_t\right) &= \mathbb{E}\left(\int_t^{t+h} \mu_s ds \middle| \mathcal{F}_t\right), \\ \operatorname{var}\left(X_{t+h} - X_t | \mathcal{F}_t\right) &= \operatorname{var}\left(\int_t^{t+h} \sigma_s dB_s \middle| \mathcal{F}_t\right) \\ &= \mathbb{E}\left(\int_t^{t+h} \sigma_s^2 ds \middle| \mathcal{F}_t\right). \end{aligned}$$

Note that X_t is a martingale with quadratic variation $\int_0^t \sigma_s^2 ds$ if and only if $\mu_t = 0$ for all t.

When μ_s and σ_s are constant over the interval $[t, t+h]$, then the conditional mean and variance reduce to $\mu_t h$ and $\sigma_t^2 h$, respectively. In general

$$\mu_t = \left.\frac{d\mathbb{E}\left(X_s | \mathcal{F}_t\right)}{ds}\right|_{s=t}, \qquad \sigma_t^2 = \left.\frac{d\operatorname{var}\left(X_s | \mathcal{F}_t\right)}{ds}\right|_{s=t},$$

the "instantaneous mean and variance" of dX_t, given $\mathcal{F}_t$.

Suppose that X_t is an Itô process with drift μ_t and volatility σ_t, and let $Y_t = f(t, X_t)$ be a function of X_t and t, with

$$\begin{aligned} f_1(t,x) &= \frac{\partial f(t,x)}{\partial t}, \qquad f_2(t,x) = \frac{\partial f(t,x)}{\partial x}, \\ f_{ij}(t,x) &= \left.\frac{\partial^2 f(x_1,x_2)}{\partial x_i \partial x_j}\right|_{(x_1,x_2)=(t,x)} \qquad i,j = 1,2. \end{aligned}$$

With ordinary differentials, we would have $df(t, X_t) = f_1(t, X_t)dt + f_2(t, X_t)dX_t$. The basis of this is that a function may be approximated locally linearly, with higher order terms of lower order than the first-order terms. With stochastic differentials, this is no longer valid, in particular because the second-order term $(dX_t)^2 = \sigma_t^2 dt$ is of the same order as the first part $\mu_t dt$ of the first-order term dX_t.

We use a second-order Taylor series expansion. Let $\Delta f(t, X_t) = f(t + \Delta t, X_{t+\Delta t}) - f(t, X_t)$ and $\Delta X_t = X_{t+\Delta t} - X_t$. Then

$$\begin{aligned} \Delta f(t, X_t) &= f_1(t, X_t)\Delta t + f_2(t, X_t)\Delta X_t \\ &\quad + \frac{1}{2}\left[f_{11}\left(\Delta t\right)^2 + 2f_{12}\Delta t \Delta X_t + f_{22}\left(\Delta X_t\right)^2\right], \end{aligned}$$

where the second-order derivatives are evaluated at (t^*, X_{t^*}), with $t^* \in [t, t+\Delta t]$.

Summing these increments, letting $\Delta t \to 0$ and using $(dX_t)^2 = \sigma_t^2 dt$, $(dt)^2 = 0$ and $dtdX_t = 0$, we obtain

$$\begin{aligned} f(t, X_t) &= f(0, X_0) + \int_0^t f_1(s, X_s)ds + \int_0^t f_2(s, X_s)dX_s \\ &\quad + \frac{1}{2}\int_0^t f_{22}(s, X_s)\sigma_s^2 ds. \end{aligned}$$

This is known as *Itô's lemma*, in integral form. In differential form it reads

$$df(t, X_t) = f_1(t, X_t)dt + f_2(t, X_t)dX_t + \frac{1}{2}f_{22}(t, X_t)\sigma_t^2 dt.$$

All this implies that $f(t, X_t)$ is also an Itô process, with drift $[f_1(t, X_t) + f_2(t, X_t)\mu_t + \frac{1}{2}f_{22}(t, X_t)\sigma_t^2]$, and volatility $f_2(t, X_t)\sigma_t$.

Example Let $f(t, X_t) = f(X_t) = \log(X_t)$. Then $f_1 = 0$, $f_2(X_t) = 1/X_t$, and $f_{22}(X_t) = -1/X_t^2$. Therefore,

$$d \log X_t = \frac{1}{X_t}dX_t - \frac{\sigma_t^2}{2X_t^2}dt.$$

This is in contrast to the usual (Newtonian) differential $d \log x = \dfrac{1}{x}dx$.

Exercises

1. Using integration by parts, show that for any differentiable function g with $g(0) = 0$, $\int_0^t g(s)dg(s) = \frac{1}{2}g(t)^2$; next, use Itô's lemma to prove $\int_0^t B_s dB_s = \frac{1}{2}(B_t^2 - t)$, by defining $f(t, B_t) = B_t^2$. Hence integration by parts in general does not hold for stochastic processes.

2. Use Itô's lemma to prove the following: if $dX_t = \alpha dt + \sigma dB_t$ and $S_t = f(t, X_t) = \exp(X_t)$, then

$$dS_t = \mu S_t dt + \sigma S_t dB_t,$$

with $\mu = \alpha + \frac{1}{2}\sigma^2$.

3. Also use Itô's lemma to prove the previous exercise "in reverse": if $dS_t = \mu S_t dt + \sigma S_t dB_t$, and $X_t = \ln(S_t)$, then $dX_t = \alpha dt + \sigma dB_t$, with $\alpha = \mu - \frac{1}{2}\sigma^2$.

4. Let $Y_t = \mu + e^{-\gamma t}[Y_0 - \mu] + \sigma \int_0^t e^{-\gamma(t-s)}dB_s$. Derive the drift and diffusion of $X_t = e^{\gamma t}Y_t$. Next, use Itô's lemma to show that $Y_t = f(t, X_t) = e^{-\gamma t}X_t$ satisfies the stochastic differential equation

$$dY_t = -\gamma(Y_t - \mu)dt + \sigma dB_t,$$

which defines it to be an Ornstein-Uhlenbeck process.

5. Suppose that we observe the process Y_t from the previous exercise at times $t = 0, h, 2h, \ldots, nh = T$. Defining $y_i = Y_{ih}$, show that

$$y_i = \mu + \rho(y_{i-1} - \mu) + \varepsilon_i, \qquad i = 1, \ldots, n,$$

where $\rho = e^{-\gamma h}$ and ε_i is an i.i.d. $N(0, \sigma_\varepsilon^2)$ sequence, with $\sigma_\varepsilon^2 = \sigma^2(1 - \rho^2)/2\gamma$.

This result implies that if $\gamma > 0$, then y_i is a mean-reverting, Gaussian first-order autoregressive sequence, which is stationary if $y_0 = Y_0$ is chosen from the stationary $N(\mu, \sigma_\varepsilon^2/(1 - \rho^2)) = N(\mu, \sigma^2/2\gamma)$ distribution. Under the same starting value assumption, the continuous-time Ornstein-Uhlenbeck process Y_t is stationary as well.