AI2613: Itô Integral, Itô Formula

May 21, 2024

1 Itô Integral

Recall that in the last lecture, we formalized a diffusion $\{X_t\}$ as

$$\mathrm{d}X_t = \mu(t, X_t)\,\mathrm{d}t + \sigma(t, X_t)\,\mathrm{d}B_t, \tag{1}$$

where $\{B_t\}$ is the standard Brownian motion. With this formalization, the motion of $\{X_t\}$ in a tiny time interval $[t, t+h]$ can be viewed as a moving under the differential equation $\frac{\mathrm{d}X_t}{\mathrm{d}t} = \mu(t, X_t)\,\mathrm{d}t$ with a random perturbation $\sigma(t, X_t)\,\mathrm{d}B_t$. In this lecture, we provide the exact mathematical meaning of the above stochastic differential equation.

Given an ordinary differential equation $\mathrm{d}f(t) = f(t)\,\mathrm{d}t$, we have that,

$$\forall T,\ \int_0^T \mathrm{d}f(t) = \int_0^T f(t)\,\mathrm{d}t,$$

which is equivalent to

$$\forall T,\ f(t) = f(0) + \int_0^T f(t)\,\mathrm{d}t.$$

If we apply the same process to Equation (1), we have that

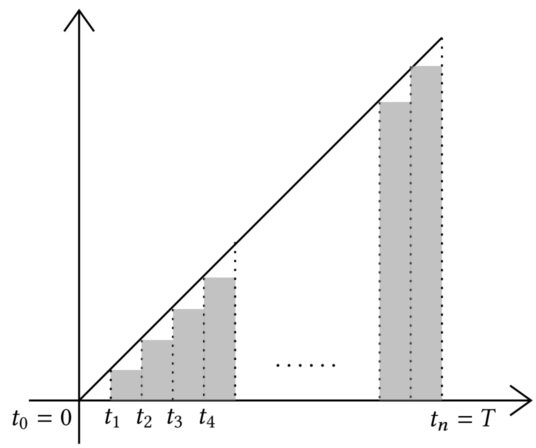
$$\forall T,\ X_T = X_0 + \int_0^T \mu(t, X_t)\,\mathrm{d}t + \int_0^T \sigma(t, X_t)\,\mathrm{d}B_t. \tag{2}$$

Let Ω be the sample space on which $\{X_t\}$ is defined. For a fixed $\omega \in \Omega$, both $\{X_t(\omega)\}$ and $\{B_t(\omega)\}$ are fixed functions with regard to t. Then Equation (2) means $\forall T, \forall \omega \in \Omega$,

$$X_t(\omega) = X_0(\omega) + \int_0^T \mu(t, X_t(\omega))\,\mathrm{d}t + \int_0^T \sigma(t, X_t(\omega))\,\mathrm{d}B_t(\omega).$$

We will omit the ω in the following if the identity holds for *almost every* ω. Note that $\int_0^T \mu(t, X_t)\,\mathrm{d}t$ is the ordinary Riemann integral of $\mu(t, X_t)$. The main goal today is to rigorously define the meaning of $\int_0^T \sigma(t, X_t)\,\mathrm{d}B_t$.

We first recall the Riemann integral and the Riemann-Stieltjes integral.



1.1 Riemann Integral

Consider the simple integral $\int_0^T x\,\mathrm{d}x$. As the following figure shows, we divide $[0, T]$ into n disjoint segments $[0, t_1], (t_1, t_2], \ldots, (t_{n-1}, T]$. Let $\Delta_i = t_i - t_{i-1}$ and $\Delta = \max_{i\in[n]} \Delta_i$. Then

$$\int_0^T x\,\mathrm{d}x = \lim_{\Delta\to 0} \sum_{i=1}^{n} t_{i-1}(t_i - t_{i-1}) = \lim_{\Delta\to 0} -\frac{1}{2}\sum_{i=1}^{n}(t_i - t_{i-1})^2 + \frac{1}{2}\sum_{i=1}^{n}(t_i^2 - t_{i-1}^2).$$

Note that $\sum_{i=1}^{n}(t_i^2 - t_{i-1}^2) = T^2$ and

$$\sum_{i=1}^{n}(t_i - t_{i-1})^2 \le \left(\max_{j\in[n]} \Delta_j\right) \cdot \sum_{i=1}^{n}(t_i - t_{i-1}) = \left(\max_{j\in[n]} \Delta_j\right) \cdot T \longrightarrow 0.$$

Therefore, we have $\int_0^T x\,\mathrm{d}x = \frac{T^2}{2}$.

1.2 Riemann-Stieltjes Integral

Then we use similar idea to calculate Riemann-Stieltjes integral[1]

Let $F\colon [0, T] \longrightarrow \mathbb{R}$ be a function with bounded derivative. Similarly, we divide $[0, T]$ into n disjoint segments $[0, t_1], (t_1, t_2], \ldots, (t_{n-1}, T]$. Assume notations the same as Section 1.1. Then the Riemann-Stieltjes

[1] The Riemann-Stieltjes integral has been widely used in probability theory. Let X be a random variable on sample space $[0, 1]$. Assume that the CDF of X is F and the PDF of X is f. Then the expectation of X is $\mathbf{E}[X] = \int_0^1 X\,\mathrm{d}F(t)$. By the definition of the Riemann-Stieltjes integral, we have

$$\int_0^1 X\,\mathrm{d}F = \lim_{n\to\infty} X(t_i^*)\,(F(t_i) - F(t_{i-1})) = \lim_{n\to\infty} X(t_i^*) f(t_{i-1})(t_i - t_{i-1}) + o(1),$$

which yields that $\mathbf{E}[X] = \int_0^1 X_t\,\mathrm{d}F(t) = \int_0^T X_t f(t)\,\mathrm{d}t$.

integral of x with respect to F is defined by

$$\begin{aligned}\int_0^T x \,\mathrm{d}F(x) &\triangleq \lim_{\Delta\to 0}\sum_{i=1}^n t_{i-1}\left(F(t_i)-F(t_{i-1})\right)\\ &= \lim_{\Delta\to 0}\sum_{i=1}^n t_{i-1}\cdot F'(t_{i-1})\left(t_i-t_{i-1}\right)\\ &= \int_0^T xF'(x)\,\mathrm{d}x\end{aligned}$$

We can use the same idea as Section 1.1 to calculate $\int_0^T F(x)\,\mathrm{d}F(x)$:

$$\begin{aligned}\int_0^T F(x)\,\mathrm{d}F(x) &= \lim_{\Delta\to 0}\sum_{i=1}^n F(t_{i-1})\left(F(t_i)-F(t_{i-1})\right)\\ &= \frac{1}{2}\left(F(T)^2-F(0)^2\right)-\lim_{\Delta\to 0}\frac{1}{2}\sum_{i=1}^n\left(F(t_i)-F(t_{i-1})\right)^2.\end{aligned}$$

If the derivative of F is bounded by M on $[0,T]$, we have

$$\begin{aligned}\lim_{\Delta\to 0}\sum_{i=1}^n\left(F(t_i)-F(t_{i-1})\right)^2 &\le \lim_{\Delta\to 0}\left(\max_{j\in[n]}\left|F(t_j)-F(t_{j-1})\right|\right)\cdot\sum_{i=1}^n\left|\left(F(t_i)-F(t_{i-1})\right)\right|\\ &\le \lim_{\Delta\to 0}\left(\max_{j\in[n]}\left|F(t_j)-F(t_{j-1})\right|\right)\cdot\sum_{i=1}^n M\cdot(t_i-t_{i-1})\\ &= \lim_{\Delta\to 0}\left(\max_{j\in[n]}\left|F(t_j)-F(t_{j-1})\right|\right)\cdot MT = 0.\end{aligned}$$

1.3 Itô Integral

Consider what will happen if we substitute F with the standard Brownian motion $\{B_t\}$. We can also deduce that

$$\int_0^T B_t\,\mathrm{d}B_t = \frac{1}{2}B_T^2-\frac{1}{2}\lim_{\Delta\to 0}\sum_{i=1}^n\left(B_{t_i}-B_{t_{i-1}}\right)^2.$$

However, as $B_t(\omega)$ could be non-differentiable with regard to t, we can not calculate this integral as we do in the Riemann-Stieltjes integral. The term $\lim_{\Delta\to 0}\sum_{i=1}^n\left(B_{t_i}-B_{t_{i-1}}\right)^2$ might not vanish! In this following, we aim at understanding what it is.

Let $Q_n=\sum_{i=1}^n\left(B_{t_i}-B_{t_{i-1}}\right)^2$. It is a random variable, and thus we first compute its expectation and variance. Recall that $\Delta_i=t_i-t_{i-1}$. We know $B_{t_i}-B_{t_{i-1}}\sim\mathcal{N}(0,\Delta_i)$. Then

$$\mathbf{E}\left[Q_n\right]=\sum_{i=1}^n\mathbf{E}\left[\left(B_{t_i}-B_{t_{i-1}}\right)^2\right]=\sum_{i=1}^n\Delta_i=T,$$

and

$$\begin{aligned}\mathbf{Var}\left[Q_n\right] &= \sum_{i=1}^{n} \mathbf{Var}\left[\left(B_{t_i} - B_{t_{i-1}}\right)^2\right] = \sum_{i=1}^{n} \mathbf{E}\left[\left(B_{t_i} - B_{t_{i-1}}\right)^4\right] - \sum_{i=1}^{n} \mathbf{E}\left[\left(B_{t_i} - B_{t_{i-1}}\right)^2\right]^2 \\ &= 2\sum_{i=1}^{n} \Delta_i^2 \leq 2\left(\max_{i\in[n]} \Delta_i\right) \cdot \sum_{i=1}^{n} \Delta_i.\end{aligned}$$

The variance tends to 0 as $\Delta \longrightarrow 0$ (or equivalently as $n \longrightarrow \infty$ if we choose each Δ_i to have equal length). This means that Q_n converges to T in the following mean square sense.

Definition 1 (Mean Square Convergence). *Let $Z_1, Z_2, \ldots$ and Z be random variables that $\mathbf{E}\left[Z^2\right] < \infty$ and $\mathbf{E}\left[Z_n^2\right] < \infty$ for $n \geq 1$. We say Z_n converges to Z in mean square, or Z is the mean square limit of $\{ Z_n \}$, if $\lim_{n\to\infty} \mathbf{E}\left[(Z_n - Z)^2\right] = 0$. This is denoted as $Z_n \xrightarrow{L_2} Z$.*

Then we can define the Itô integral.

Definition 2. *Assume that $\{X_t\}$ is a "nice enough" stochastic process* [2]*. Then we define the integral $\int_0^T X_t \, \mathrm{d}B_t$ as the mean square limit of*

$$\sum_{i=1}^{n} X_{t_{i-1}} \left(B_{t_i} - B_{t_{i-1}}\right).$$

This is called the Itô integral of $\{ X_t \}$ with respect to $\{ B_t \}$.

With Definition 2, we can verify that

$$\int_0^T B_t \, \mathrm{d}B_t = \frac{1}{2}B_T^2 - \frac{1}{2}T.$$

More generally, we may define $\int_0^T X_t \, \mathrm{d}B_t$ as the mean square limit of

$$\sum_{i=1}^{n} X_{t_i^*} \left(B_{t_i} - B_{t_{i-1}}\right),$$

where $t_i^* = \alpha t_{i-1} + (1-\alpha)t_i$ with $\alpha \in [0, 1]$. The Itô integral defined in Definition 2 corresponds to the case that $\alpha = 1$. By choosing $\alpha = \frac{1}{2}$, we have the definition of Stratonovich integral and it holds that $\int_0^T B_t \, \mathrm{d}B_t = \frac{1}{2}B_T^2$ with Stratonovich integral.

2 Itô Formula

Recall that in the example in Section 1.3, we have

$$Q_n = \sum_{i=1}^{n} \left(B_{t_i} - B_{t_{i-1}}\right)^2, \quad \mathbf{E}\left[Q_n\right] = T \quad \text{and} \quad \mathbf{Var}\left[Q_n\right] \longrightarrow 0.$$

[2] This means that the process has some nice properties, such as that X_t is independent with $\{B_u\}_{u>t}$ for all t. Some other technical requirements can be found in any standard textbook on stochastic differential equations (e.g., [Kle12])

In other words,

$$\lim_{n\to\infty} Q_n = \int_0^T (\mathrm{d}B_t)^2 = T = \int_0^T \mathrm{d}t.$$

This suggests that $(\mathrm{d}B_t)^2 \approx \mathrm{d}t$ holds. By definition of $\{ B_t \}$, we know that $\mathrm{d}B_t = B_{t+\mathrm{d}t} - B_t \sim \mathcal{N}(0, \mathrm{d}t)$. Let X and Y be two random variables that $X \sim \mathcal{N}(0, \mathrm{d}t)$ and $Y = X^2$. Then the formula $(\mathrm{d}B_t)^2 \approx \mathrm{d}t$ tells us that Y is well concentrated on $\mathrm{d}t$.

With this observation, we then (heuristically) deduce the chain rule under the definition of Itô integral.

2.1 Classical Chain Rule of Differentiation

For any differentiable function f, we have $\mathrm{d}f(t) = f'(t)\,\mathrm{d}t$. This can be verified via Taylor's expansion:

$$\mathrm{d}f(t) = f(t+\mathrm{d}t) - f(t) = f'(t)\,\mathrm{d}t + \frac{1}{2}f''(t)\,(\mathrm{d}t)^2 + o\left((\mathrm{d}t)^2\right) \overset{\mathrm{d}t\to 0}{\longrightarrow} f'(t)\,\mathrm{d}t.$$

For two differentiable functions g and f, the chain rule of differentiation is

$$\frac{\mathrm{d}f\left(g(t)\right)}{\mathrm{d}t} = f'\left(g(t)\right)\cdot g'(t).$$

We can also derive this using the Taylor expansion:

$$\begin{aligned}\mathrm{d}f\left(g(t)\right) &= f\left(g(t+\mathrm{d}t)\right) - f\left(g(t)\right)\\ &= f\left(g(t) + \mathrm{d}g(t)\right) - f\left(g(t)\right)\\ &= f'\left(g(t)\right)\mathrm{d}g(t) + \frac{1}{2}f''\left(g(t)\right)\left(\mathrm{d}g(t)\right)^2 + o((\mathrm{d}g(t))^2).\end{aligned}$$

Then it follows that

$$\frac{\mathrm{d}f\left(g(t)\right)}{\mathrm{d}t} = f'\left(g(t)\right)\cdot g'(t) + o(\mathrm{d}t) \overset{\mathrm{d}t\to 0}{\longrightarrow} f'\left(g(t)\right)\cdot g'(t).$$

2.2 The Chain Rule with Itô Integral

Consider a diffusion $\{X_t\}$ such that

$$\mathrm{d}X_t = \mu_t\,\mathrm{d}t + \sigma_t\,\mathrm{d}B_t,$$

where μ_t and σ_t are the abbreviations of $\mu(t, X_t)$ and $\sigma(t, X_t)$ respectively. By the Taylor's expansion,

$$\begin{aligned}\mathrm{d}f\left(X_t\right) &= f\left(X_t + \mathrm{d}X_t\right) - f\left(X_t\right)\\ &= f'\left(X_t\right)\mathrm{d}X_t + \frac{1}{2}f''\left(X_t\right)\left(\mathrm{d}X_t\right)^2 + o\left((\mathrm{d}X_t)^2\right)\\ &= f'\left(X_t\right)\left(\mu_t\,\mathrm{d}t + \sigma_t\,\mathrm{d}B_t\right) + \frac{1}{2}f''\left(X_t\right)\left(\mu_t\,\mathrm{d}t + \sigma_t\,\mathrm{d}B_t\right)^2 + o\left((\mathrm{d}X_t)^2\right)\\ &\overset{\mathrm{d}t\to 0}{\longrightarrow} \left(f'\left(X_t\right)\mu_t + \frac{1}{2}f''\left(X_t\right)\sigma_t^2\right)\mathrm{d}t + f'\left(X_t\right)\sigma_t\,\mathrm{d}B_t\end{aligned}$$

This chain rule with Itô integral is called Itô formula. Then we see some examples of Itô formula.

Example 1. *We can use Itô formula to calculate* B_T^2*. Note that*

$$\mathrm{d}\left(B_t\right)^2 = 2B_t\,\mathrm{d}B_t + \left(\mathrm{d}B_t\right)^2 = 2B_t\,\mathrm{d}B_t + \mathrm{d}t.$$

Therefore, $B_T^2 = 2\int B_t\,\mathrm{d}B_t + T$*. This provides a quick justification that* $\int_0^T B_t\,\mathrm{d}B_t = \frac{1}{2}B_t^2 - \frac{1}{2}T$*.*

Example 2 (Geometric Brownian Motion)**.** *The geometric Brownian motion is* $Y_t = e^{B_t}$ *where* $\{B_t\}$ *is the standard Brownian motion. Then it follows from the Itô formula that*

$$\mathrm{d}Y_t = e^{B_t}\left(\mathrm{d}B_t + \frac{1}{2}\left(\mathrm{d}B_t\right)^2\right) = Y_t\,\mathrm{d}B_t + \frac{1}{2}Y_t\,\mathrm{d}t.$$

Example 3 (Ornstein-Uhlenbeck Process)**.** *Let* $\{X_t\}$ *be a Ornstein-Uhlenbeck process that* $\mathrm{d}X_t = -X_t\,\mathrm{d}t + 2\,\mathrm{d}B_t$ *with* $X_0 = 0$*. We can calculate* X_t *according to this equation. Let* $f(t, X_t) = e^t \cdot X_t$*. Using Taylor expansion, we have*

$$\begin{aligned}
\mathrm{d}f(t, X_t) &= f(t + \mathrm{d}t, X_{t+\mathrm{d}t}) - f(t, X_t) \\
&= X_t e^t\,\mathrm{d}t + e^t\,\mathrm{d}X_t \\
&= X_t e^t\,\mathrm{d}t - e^t \cdot X_t\,\mathrm{d}t + 2e^t\,\mathrm{d}B_t = 2e^t\,\mathrm{d}B_t.
\end{aligned}$$

Then we have $e^T X_T = \int_0^T 2e^t\,\mathrm{d}B_t$ *and* $X_t = e^{-T}\int_0^T 2e^t\,\mathrm{d}B_t$*. Since* $\int_0^T 2e^t\,\mathrm{d}B_t$ *can be viewed as the limit of the weighted sum of many Gaussian random variables, one can verify that* $\int_0^T 2e^t\,\mathrm{d}B_t \sim \mathcal{N}\left(0, 4\int_0^T e^{2t}\,\mathrm{d}t\right)$*.*

References

[Kle12] Fima C Klebaner. *Introduction to stochastic calculus with applications.* World Scientific Publishing Company, 2012. 4