2 Stochastic integration

Central for stochastic calculus is the concept of Itô's integral

$$\int_0^T X(t)dB(t), \tag{7}$$

whose basic ingredients are the Brownian motion with filtration $(\mathcal{F}_t, t \geq 0)$ (which models the information flow) and a stochastic process $(X_t, t \geq 0)$ adapted to the filtration. Although construction of the Itô integral is similar to that from the classical calculus, the Itô integral is different due to the nontrivial quadratic variation of BM.

2.1 Introductory example

For continuously differentiable function f with $f(0) = 0$ the integral $\int_0^T f(t)df(t)$ is calculated as

$$\int_0^T f(t)df(t) = \int_0^T f(t)f'(t)dt = \frac{1}{2}f^2\Big|_0^T = \frac{1}{2}f^2(T).$$

This integral can be defined as a limit of Riemann-Stiltjes integral sums

$$\sum_{i=0}^{n-1} f(t_i)(f(t_{i+1}) - f(t_i))$$

as the mesh size $|\varDelta| \to 0$.

Following this line, let us see what happens if we define the integral

$$\int_0^T B(t)dB(t).$$

For simplicity we take the uniform partition of $[0, T]$ by points $t_i = Ti/n$, and we write $B_i := B(t_i)$ for the values of BM at these points. With some algebra we have

$$\frac{1}{2}\sum_{i=0}^{n-1}(B_{i+1} - B_i)^2 = \frac{1}{2}\sum_{i=0}^{n-1} B_{i+1}^2 - \sum_{i=0}^{n-1} B_i B_{i+1} + \frac{1}{2}\sum_{i=0}^{n-1} B_i^2 =$$

$$\frac{1}{2}B_n^2 + \frac{1}{2}\sum_{i=0}^{n-1} B_i^2 - \sum_{i=0}^{n-1} B_i B_{i+1} + \frac{1}{2}\sum_{i=0}^{n-1} B_i^2 = \frac{1}{2}B_n^2 + \sum_{i=0}^{n-1} B_i^2 - \sum_{i=0}^{n-1} B_i B_{i+1} =$$

$$\frac{1}{2}B_n^2 - \sum_{i=0}^{n-1} B_i(B_{i+1} - B_i),$$

which yields

$$\sum_{i=0}^{n-1} B_i(B_{i+1} - B_i) = \frac{1}{2}B_n^2 - \frac{1}{2}\sum_{i=0}^{n-1}(B_{i+1} - B_i)^2.$$

That is, the integral sum can be written as

$$\sum_{i=0}^{n-1} B(Ti/n)(B(T(i+1)/n) - B(Ti/n)) =$$
$$\frac{1}{2}B(T)^2 - \frac{1}{2}\sum_{i=0}^{n-1}(B(T(i+1)/n) - B(Ti/n))^2.$$

Letting $n \to \infty$, and understanding the limit as in Theorem 1.13

$$\sum_{i=0}^{n-1}(B(T(i+1)/n) - B(Ti/n))^2 \to \langle B \rangle(T) = T,$$

which is the quadratic variation (6). Therefore

$$\sum_{i=0}^{n-1} B(Ti/n)(B(T(i+1)/n) - B(Ti/n)) \to B^2(T) - \frac{1}{2}T.$$

If B *were* a smooth function we would get $\frac{1}{2}B(T)$. The extra term $\frac{1}{2}T$ comes from the quadratic variation of BM.

2.2 Simple integrands

Like the classical integrals, the integral (7) is first defined for simple functions, then extended to a larger class.

Start with the simplest case $X(t) = 1(t \in [a, b))$ of (nonrandom) indicator function of interval. For the integral – just to justify its name! – we certainly want to have

$$\int_a^b dB(t) = B(b) - B(a).$$

Next by complexity case is $X(t) = \xi\, 1(t \in [a, b))$ with some $\mathcal{F}_a$-measurable random variable ξ. Naturally, we want to have the linearity, so we define the integral as

$$\int_a^b \xi dB(t) = \xi\, (B(b) - B(a)). \tag{8}$$

Note that $(X(t), t \geq 0)$ is an adapted process due to $\mathcal{F}_a$-measurability of ξ. Trivially, $X(t) = 0$ for $t \notin [a, b)$, and for $t \in [a, b)$ we have $X(t) = \xi$ hence $\mathcal{F}_t$-measurable, since $\mathcal{F}_t \supset \mathcal{F}_a$ (information accumulates). Furthermore, with the measurability condition we achieve that the integral

$$I(t) = \int_0^t \xi\, 1(u \in [a, b))dB(u),$$

considered as a function of the upper limit of integration t, is a martingale. Let us check the martingale condition $\mathbb{E}(I(t)|\mathcal{F}_s) = I(s)$ for $a \leq s < t \leq b$. Using (8) with upper limit t we compute

$$\mathbb{E}[I(t)|\mathcal{F}_s] = \mathbb{E}[\xi(B(t) - B(a))|\mathcal{F}_s] = \xi\mathbb{E}[B(t) - B(a)|\mathcal{F}_s] =$$
$$\xi\{\mathbb{E}[B(t)|\mathcal{F}_s] - \mathbb{E}[B(a)|\mathcal{F}_s]\} = \xi(B(s) - B(a)) = I(s),$$

because ξ and $B(a)$ are $\mathcal{F}_s$-measurable and $\mathbb{E}[B(t)|\mathcal{F}_s] = B(s)$ (BM is a martingale).

Let $0 = t_0 < t_1 < \cdots < t_n = T$ be points dividing $[0, T]$ is n subintervals $[t_i, t_{i+1})$, and let ξ_i be a $\mathcal{F}_{t_i}$-measurable random variable with $\mathbb{E}\xi_i^2 < \infty$ for $i = 0, \ldots, n-1$. The process

$$X(t) = \sum_{i=0}^{n-1} \xi_i 1(t \in [t_i, t_{i+1}))$$

is called *simple*. The simple process is

- adapted to $(\mathcal{F}_t, t \geq 0)$,
- piecewise-constant as function of t: $X(t) = X(t_i) = \xi_i$ for $t \in [t_i, t_{i+1})$ and $0 \leq i \leq n-1$, and $X(t) = 0$ for $t \geq T$.

Definition 2.1. The stochastic integral (the Itô integral) over $[0, T]$ for simple process is defined as

$$I(T) = \int_0^T X(t)dB(t) = \sum_{i=0}^{n-1} \xi_i \left(B(t_{i+1} - B(t_i)\right).$$

Similarly, the stochastic integral as a function of the upper limit is defined for $t \in [t_k, t_{k+1})$, $0 \leq k \leq n-1$ as

$$I(t) = \sum_{i=0}^{k-1} \xi_i \left(B(t_{i+1} - B(t_i)\right) + \xi_k \left(B(t) - B(t_k)\right)$$

From the definition easily follows:

- $(I(t), t \geq 0)$ is an adapted process.

Theorem 2.2. *The stochastic integral $(I(t), t \geq 0)$ of a simple process is a martingale. In particular, $\mathbb{E}I(t) = 0$.*

Proof. Fix $s < t$. We need to show that $\mathbb{E}[I(t)|\mathcal{F}_s] = I(s)$. We can always treat t and s as division points: if $t \in [t_k, t_{k+1})$ just replace $\xi_k 1(u \in [t_k, t_{k+1}))$ in $X(u)$ by two terms $\xi_k 1(u \in [t_k, t)) + \xi_k 1(u \in [t, t_{k+1})$, and do similarly for s. Thus suppose $s = t_i < t_j = t$. We have

$$I(t_j) = I(t_i) + \sum_{k=i}^{j-1} \xi_k (B(t_{k+1}) - B(t_k)).$$

Clearly, $\mathbb{E}(I(t_i)|\mathcal{F}_{t_i}) = I(t_i)$, since $I(t_i)$ is $\mathcal{F}_{t_i}$-measurable. We claim that

$$\mathbb{E}[\xi_k(B(t_{k+1}) - B(t_k))|\mathcal{F}_{t_i}] = 0, \quad k \geq i.$$

Indeed, since $\mathcal{F}_{t_i} \subset \mathcal{F}_{t_k}$ and ξ_k is $\mathcal{F}_{t_k}$-measurable

$$\mathbb{E}[\xi_k(B(t_{k+1}) - B(t_k))|\mathcal{F}_{t_i}] = \mathbb{E}\{\mathbb{E}[\xi_k(B(t_{k+1}) - B(t_k))|\mathcal{F}_{t_k}]|\mathcal{F}_{t_i}\} = \mathbb{E}\{\xi_k\mathbb{E}[B(t_{k+1}) - B(t_k)|\mathcal{F}_{t_k}]|\mathcal{F}_{t_i}\} = \mathbb{E}\{\xi_k \cdot 0|\mathcal{F}_{t_i}\} = 0.$$

Adding up all calculated terms yields $\mathbb{E}[I(t_j)|\mathcal{F}_{t_i}) = I(t_i)$, as wanted. □

From $\mathbb{E}I(t) = 0$ follows that $\mathrm{Var}I(t) = \mathbb{E}I^2(t)$.

Theorem 2.3. (Itô isometry)

$$\mathbb{E}I^2(t) = \mathbb{E}\int_0^t X^2(u)du. \tag{9}$$

Proof. Treating $t \in [0, T]$ as a division point, we reduce to the case $t = T$ $I(T)$ is a sum of terms $\xi_i(B(t_{i+1}) - B(t_i))$. The squared sum has squares $\xi_i^2(B(t_{i+1}) - B(t_i))^2$ and the cross terms $2\xi_i(B(t_{i+1}) - B(t))\xi_j(B(t_{j+1}) - B(t_j))$.

In a cross term $B(t_{j+1}) - B(t_j)$ is independent of the other factors, hence

$$\mathbb{E}\{\xi_i(B(t_{i+1}) - B(t))\xi_j(B(t_{j+1}) - B(t_j))\} = \mathbb{E}\{\xi_i(B(t_{i+1}) - B(t))\xi_j\}\mathbb{E}\{B(t_{j+1}) - B(t_j)\} = 0.$$

In a squared term $B(t_{i+1}) - B(t_i)$ is independent of ξ_i, hence

$$\mathbb{E}\{\xi_i^2(B(t_{i+1}) - B(t_i))^2\} = \mathbb{E}\xi_i^2\mathbb{E}(B(t_{i+1}) - B(t_i))^2 = (t_{i+1} - t_i)\mathbb{E}\xi_i^2.$$

Adding up,

$$\mathbb{E}I^2(T) = \sum_{i=1}^{n-1}(t_{i+1} - t_i)\mathbb{E}\xi_i^2.$$

On the other hand, $X(u) = \sum_{i=0}^{n-1} \xi_i 1(u \in [t_i, t_{i+1}))$ for each u has only one non-vanishing term, so squaring yields $X^2(u) = \sum_{i=0}^{n-1} \xi_i^2 1(u \in [t_i, t_{i+1}))$. Integrating the piecewise-constant function

$$\int_0^T X^2(u)du = \sum_{i=0}^{n-1} \xi_i^2(t_{i+1} - t_i),$$

and finally

$$\mathbb{E}\int_0^T X^2(u)du = \sum_{i=0}^{n-1}(t_{i+1} - t_i)\mathbb{E}\xi_i^2,$$

which coincides with $\mathbb{E}I^2(T)$ computed above. □

Theorem 2.4. *The quadratic variation accumulated by the Itô integral $I(t) = \int_0^t X(u)dB(u)$ up to time t is*

$$\langle I\rangle(t) = \int_0^t X^2(u)du.$$

Proof. We first look at the contribution of one partition subinterval $[t_i, t_{i+1})$ of $[0, t]$, where $X(u) = \xi_i$ $(u \in [t_i, t_{i+1})$ is a contant. Splitting the interval by some number of points s_j, then letting $\max_j |s_{j+1} - s_j| \to 0$

$$\sum_i (I(s_{j+1}) - I(s_j))^2 = \xi_i^2 \sum_j (B(s_{j+1} - B(s_j))^2 \to \xi_i^2 (t_{i+1} - t_i)$$

by the formula for the quadratic variation of BM. Summing over the 'steps' of X yields

$$\langle I \rangle (t) = \sum_i \xi_i^2 (t_{i+1} - t_i) = \int_0^t X^2(u) du.$$

□

The heuristic 'differential form' for $I(t) = \int_0^t X(u)$ is $dI(t) = X(t)dB(t)$. Analogous to the Brownian rule $(dB(t))^2 = dt$ is the formula

$$(dI(t))^2 = X^2(t)dt,$$

whose precise meaning is revealed by Theorem 2.4: the Itô integral accumulates the quadratic variation at unit rate.

2.3 General integrands

Let $X = (X(t), t \in [0, T])$ be a process adapted to the filtration $(\mathcal{F}(t), t \geq 0)$ and such that

$$\mathbb{E} \int_0^T X^2(t) dt < \infty. \tag{10}$$

As a function of two variables $(t, \omega) \in [0, T] \times \Omega$, the process X belongs to the Hilbert space of functions $L^2([0, T] \times \Omega, dt \times d\mathbb{P})$ with the norm defined by (10). Formula (9) says that the correspondence $X \to I(T)$ is an *isometry* into the Hilbert space $L^2(\Omega, d\mathbb{P})$ of square-integrable random variables.

In $L^2([0, T] \times \Omega, dt \times d\mathbb{P})$ it is possible to approximate any X by simple processes:

Lemma 2.5. *For every adapted process X on $[0, T]$ satisfying (10) there exists a sequence of simple processes $X_n = (X_n(t), t \in [0, T)$ such that*

$$\lim_{n \to \infty} \mathbb{E} \int_0^T |X_n(t) - X(t)|^2 dt = 0.$$

For X satisfying (10) we may choose an approximating sequence X_n of simple processes (as in the lemma), for which the stochastic integrals has been introduced in Definition 2.1. We define the stochastic integral of X as a limit

$$\int_0^T X(t) dB(t) := \lim_{n \to \infty} \int_0^T X_n(t) dB(t).$$

The detailed interpretation of this relation is that the random variables $I_n := \int_0^T X_n(t) dB(t)$ in the LHS converge to the random variable I in the RHS in the mean-square sense

$\mathbb{E}(I - I_n)^2 \to 0$. The limit I exists, because $X_n \to X$ implies that I_n is (by the isometry) a Cauchy sequence. The limit does not depend on the approximating sequence X_n, as is easily seen using the triangle inequality.

The integral for the general processes has the properties which we encountered when discussing the simple processes:

- $I(t)$ is continuous in t,
- $I(t)$ is $\mathcal{F}_t$-measurable,
- the integral is linear, i.e. $\int_0^t (X(t) + Y(t))dB(t) = \int_0^t X(t)dB(t) + \int_0^t Y(t)dB(t)$,
- $(I(t), t \geq 0)$ is a martingale, with $\mathbb{E}I(t) = 0, \mathrm{Var}I(t) = \mathbb{E}\int_0^t X^2(u)du$,
- $\langle I\rangle(t) = \int_0^t X^2(u)(du)$.