Itô Integrals

Vladimir Feinberg

August 22, 2021

1 Construction of the Itô integral

We recall our original objective, which is to define some stochastic notion of differential equations parameterized by b, σ to identify some process X_t satisfying

$$X'(t) = b(t, X_t) + \sigma(t, X_t) \cdot \text{"noise"} .$$

It's sensible to consider the forward Euler method, for some shrinking split sequence t_i, which would set

$$\Delta_k X = b(t_k, X_{t_k})\Delta_k t + \sigma(t_k, X_{t_k})\Delta_k V ,$$

where $\Delta_k f = f(t_{k+1}) - f(t_k)$ and V_t is some stochastic process.

Based on natural assumptions about $\Delta_k V$ such as stationarity, independent increments, and 0-mean, we might be tempted to choose Brownian motion $V_t = B_t$. In the limit as $\Delta_k t \to 0$, we would be interested in finding a definition of an integral which is consistent with the limit of the discrete sums implied by the Euler method:

$$X_t = X_0 + \int_0^t b(s, X_s)\,\mathrm{d}s + \int_0^t \sigma(s, X_s)\,\mathrm{d}B_s ,$$

where $\int_0^t \sigma(s, X_s)\,\mathrm{d}B_s$ is not yet sensibly defined or shown to exist uniquely as $\lim_n \sum_{k=1}^n \sigma(t_k, X_{t_k})\Delta_i B$.

Definition 1.1 (Itô Integral). Consider the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ where $\mathcal{F} = \mathcal{F}_\infty$ for a filtration $\{\mathcal{F}_t\}$. Let $\mathcal{V} = \mathcal{V}(S, T)$ be the class of functions $f : \mathbb{R}_+ \times \Omega \to \mathbb{R}$ satisfying the following properties:

1. f is $\mathcal{B}(\mathbb{R}_+) \times \mathcal{F}$ measurable.

2. $f(t, \omega)$ is $\mathcal{F}_t$-adapted.

3. $L^2(\mathbb{P} \times [S, T]) = \mathbb{E}^{1/2}\left[\int_S^T f(t, \omega)^2\,\mathrm{d}t\right] < \infty$

Then under some shrinking split sequence (see previous chapter notes), the Itô Integral is defined as the limit

$$\int_S^T f(t, \omega)\,\mathrm{d}B_t(\omega) = \lim_{n \to \infty} \sum_{k=1}^n \phi_n(t_k, B_{t_k}(\omega))\Delta_i B(\omega) ,$$

where ϕ_n are simple functions chosen by diagonalizing multiple other unique limit sequences. These ϕ_n will be recovered from multiple other sequences as we grow the complexity of f to the full extent of $\mathcal{V}$. To start, if f is a simple function, then of course $\phi_n = f$ suffices. Afterwords, we show the convergence under $L^2(\mathbb{P} \times [S, T])$ of limit sequences for growing subsets of $\mathcal{V}$:

1. Bounded and everywhere continuous processes. Since simple functions approximate our continuous process everywhere on Ω, by bounded convergence they approximate in $L^2(\mathbb{P} \times [S, T])$.

2. Bounded processes $g \in \mathcal{V}$. As $n \to \infty$, consider a bump function ψ_n with shrinking support immediately left of the origin. Consider the convolution $g_n = \psi_n * g$, also in $\mathcal{V}$ by virtue of the negative support of ψ_n. Note that $g_n \to g$ pointwise as ψ_n progressively approximates the Dirac measure. Then again by bounded convergence we have $g_n \to g$ in $L^2(\mathbb{P} \times [S, T])$, where g_n are bounded and everywhere continuous by construction.

3. Any process in $g \in \mathcal{V}$. Let g_n be g clamped to $[-n, n]$. By the previous step, we have convergence to each such g_n via limit sequence in $\mathcal{V}$. Then since $g \in L^2(\mathbb{P} \times [S, T])$ itself, by dominated convergence we have convergence in $g \in L^2(\mathbb{P} \times [S, T])$.

For any $f \in \mathcal{V}$ we may diagonalize the above sequences for some *elementary* functions ϕ_n to define the Itô integral. Why does this definition give rise to a *unique* integral, independent of the choice of ϕ_n? To see this, we need a critical tool.

Lemma 1 (Itô Isometry). For any $f \in \mathcal{V}$,

$$\mathbb{E}\left[\int_S^T f(t, \omega)^2 \, \mathrm{d}t \right] = \mathbb{E}\left[\left(\int_S^T f(t, \omega) \, \mathrm{d}t \right)^2 \right] ,$$

or, in operator form,

$$L^2(\mathbb{P} \times [S, T])(f) = L^2(\mathbb{P}) \circ \int_S^T \mathrm{d}t .$$

For elementary functions, this can be proven by computing both sides explicitly and relying on stationarity of increments. In turn, this implies that for any $f \in \mathcal{V}$, for any $L^2(\mathbb{P} \times [S, T])$ sequence ϕ_n of elementary functions, which is necessary Cauchy in that metric, must give rise to a sequence of integrals $I_n = \int_S^T \phi_n \, \mathrm{d}B_t$ which are Cauchy in $L^2(\mathbb{P})$ and thus converge as random variables in this way to a fixed limit $I = I(f, \phi)$ in $L^2(\mathbb{P})$. But for any other similar elementary function sequence ϕ_n' converging to f in $L^2(\mathbb{P} \times [S, T])$, we construct an analogous sequence of integrals I_n', also Cauchy in $L^2(\mathbb{P})$, thus also yielding an Itô integral $I' = I(f, \phi')$. Through our elementary isometry,

$$\|I_n - I'\|_{L^2(\mathbb{P})} \leq \|I_n - I_n'\|_{L^2(\mathbb{P})} + \|I_n' - I'\|_{L^2(\mathbb{P})}$$
$$= \|\phi_n - \phi_n'\|_{L^2(\mathbb{P}) \circ \int_S^T \mathrm{d}t} + \|I_n' - I'\|_{L^2(\mathbb{P})}$$
$$= \|\phi_n - \phi_n'\|_{L^2(\mathbb{P} \times [S, T])} + \|I_n' - I'\|_{L^2(\mathbb{P})}$$
$$\to 0 ,$$

but since I_n converges to both I, I' in $L^2(\mathbb{P})$, by completeness of the corresponding metric space, $I = I'$ almost surely, showing both invariance of the Itô integral to the specific limit sequence ϕ_n chosen and Itô isometry on all of $\mathcal{V}$.

2 Some properties of the Itô integral

By taking limits from the definition, the following properties of the Itô integral hold almost always:

1. $\int_S^T f \, \mathrm{d}B_t = \int_S^U f \, \mathrm{d}B_t + \int_U^T f \, \mathrm{d}B_t$,

2. $\int_S^T (cf + g) \, \mathrm{d}B_t = c \int_S^T f \, \mathrm{d}B_t + \int_S^T g \, \mathrm{d}B_t$ for constant c,

3. $\mathbb{E} \int_S^T f \, \mathrm{d}B_t = 0$,

4. $\int_S^T f \, \mathrm{d}B_t \in \mathcal{F}_T$,

5. and $M_s = \int_0^s f \, \mathrm{d}B_t$ is a martingale.

Theorem 1. *Let $f \in \mathcal{V}(0,T)$. Then there exists a t-continuous version of $I_t(\omega) = \int_0^t f(s,\omega) \, \mathrm{d}B_s(\omega)$.*

Proof. Let ϕ_n be elementary functions tending to f in $L^2(\mathbb{P} \times [0,T])$ from Itô's construction, with corresponding Itô integrals $I_t^{(n)}$. For any n,m, $I_t^{(n)} - I_t^{(m)}$ is a martingale since its components are. Doob's Martingale Inequality, which is basically Markov for martingales, relates

$$\mathbb{P} \left\{ \sup_{0 \le t \le T} \left| I_t^{(n)} - I_t^{(m)} \right| \ge \epsilon \right\} \le \frac{1}{\epsilon^2} \mathbb{E} \left| I_t^{(n)} - I_t^{(m)} \right|^2 = \frac{1}{\epsilon^2} \| \phi_n - \phi_m \|_{L^2(\mathbb{P} \times [0,T])}^2 .$$

Since the RHS vanishes as n,m grow large, a subsequence $I_t^{(n_k)}$ may be chosen ensuring with geometrically decreasing ϵ between adjacent $I_t^{(n_k)}, I_t^{(n_{k+1})}$. This allows Borel-Cantelli to be applied over the set of events $\left\{ \sup_{0 \le t \le T} \left| I_t^{(n_k)} - I_t^{(n_{k+1})} \right| \ge 2^{-k} \right\}$ for varying $k \in \mathbb{N}$, so in fact with probability one $\sup_{0 \le t \le T} \left| I_t^{(n_k)} - I_t^{(n_{k+1})} \right| \le 2^{-k}$ so long as $k \ge K(\omega)$. This provides uniform convergence for almost all ω, and thus continuity in the limit. $\qquad\square$

3 Extensions of the Itô integral

Without much effort, the filtration $\mathcal{F}$ can be expanded to any other filtration $\mathcal{H}$ such that $\mathcal{F}_t \subset \mathcal{H}_t$.

Relaxing the integrator to be an arbitrary semimartingale instead of Browninan motion is possible. Revisiting our construction steps, the only direct properties used were bounded quadratic variation and adaptedness. These can be further relaxed to just semimartingale properties.

The second moment condition in Itô's construction can be relaxed to a mere condition requiring that the second moment of our integrand is almost surely finite, $\mathbb{P} \left\{ \int_S^T f(s,\omega)^2 \, \mathrm{d}s < \infty \right\} = 1$, however, this of course will end up creating non-integrable integrals, and thus drop the martingale property. However, continuity can be recovered, along with a local martingale property.

There also exists another interpretation of stochastic integrals. The Stratonovich is built from a trapezoidal rule (unlike Itô's left-sided integration rule) as the building block:

$$\int_S^T f(t,\omega) \circ \mathrm{d}B_t(\omega) = \lim_{n \to \infty} \sum_{k=1}^n \phi_n \left(\frac{t_k + t_{k+1}}{2}, B_{t_k}(\omega) \right) \Delta_i B(\omega) .$$

One reason the Stratonovich interpretation might be more natural is because of the following motivation. If we consider t-continuously differentiable approximations $B_t^{(n)}$ to our noise process B_t which on bounded intervals converge uniformly almost always, then the solution $X_t^{(n)}$ to the differential equation (for constant ω)

$$\frac{dX_t}{dt} = b(t, X_t) + \sigma(t, X_t) \frac{dB_t^{(n)}}{dt} ,$$

converges to the Stratonovich solution $X_t = X_0 + \int_0^t b(t, X_s) \, \mathrm{d}s + \int_0^t \sigma(t, X_s) \circ \mathrm{d}B_s$, whereas the solution expressed in the Itô interpretation would be

$$X_t = X_0 + \int_0^t b(t, X_s) \, \mathrm{d}s + \frac{1}{2} \int_0^t \sigma'(s, X_s) \sigma(s, X_s) \, \mathrm{d}s + \int_0^t \sigma(t, X_s) \, \mathrm{d}B_s$$

4 Exercises

4.1

We show

$$\int_0^t s \, \mathrm{d}B_s = tB_t - \int_0^t B_s \, \mathrm{d}s$$

directly from the definition of Itô integrals. We consider the progressively finer elementary functions defined by the discretization of the identity integrand as $n \to \infty$ for some shrinking split sequence t_j, which tend to the identity in $L^2(\mathbb{P} \times [0, t])$.

$$\sum_i t_i \left(B_{t_{i+1}} - B_{t_i} \right) = t_n B_t - \sum_i \left(t_{i+1} - t_i \right) B_{t_i} \ ,$$

so that the limit converges to the desired solution, because almost-always continuity of B_t permits ω-pointwise Riemann integration.

4.2

We show

$$\int_0^t B_s^2 \, \mathrm{d}B_s = \frac{1}{3} B_t^3 - \int_0^t B_s \, \mathrm{d}s$$

directly from the definition of Itô integrals. A more general technique for solving such integrals from the definition follows in three steps:

1. Rewrite the expected naive Riemann integral solution (here, $\frac{1}{3}B_t^3$) as a telescoping sum. $B_t^3 = \sum_j \Delta_j B^3$.

2. Express each term in an Itô-friendly way, i.e., as functions of increments or the left sided-timestep, i.e., $B_{t_j}, \Delta_j B$.

$$\begin{aligned}
\Delta_j B^3 &= \Delta_j B \left(B_{t_{j+1}}^2 + B_{t_{j+1}} B_{t_j} + B_j^2 \right) \\
&= \Delta_j B \left(B_{t_{j+1}}^2 - 2B_{t_{j+1}} B_{t_j} + B_j^2 \right) + 3 B_{t_{j+1}} B_{t_j} \Delta_j B \\
&= \Delta_j^3 B + 3(\Delta_j B + B_{t_j}) B_{t_j} \Delta_j B \\
&= \Delta_j^3 B + 3 B_{t_j} \Delta_j^2 B + 3 B_{t_j}^2 \Delta_j B
\end{aligned}$$

3. Summing the above formula over j and rearranging terms from the first telescope we have

$$\sum_j B_{t_j}^2 \Delta_j B = \frac{1}{3} B_t^3 - \frac{1}{3} \sum_j \Delta_j^3 B - \sum_j B_{t_j} \Delta_j^2 B \ .$$

In the limit as $n \to \infty$, the LHS converges to the desired Itô integral $\int_0^t B_s^2 \, \mathrm{d}B_s$ in $L^2(\mathbb{P})$. The middle term vanishes since $\left| \sum_j \Delta_j^3 B \right| \leq \sum_j \left| \Delta_j^3 B \right| \to [B]_t^{(3)} = 0$. To finish, it remains to show that the last term simplifies to $\int_0^t B_s \, \mathrm{d}s$.

Notice it suffices to show that $\| \sum_j B_{t_j} \Delta_j^2 B - \sum_j B_{t_j} \Delta_j t \|_{L^2(\mathbb{P})} \to 0$ as the L^2 convergence implies an almost surely convergent subsequence, and by uniqueness of the Itô integral definition, this shows almost sure equality on the final integrals regardless of the elementary function sequence we look at. Indeed,

$$\mathbb{E} \left(\sum_j B_{t_j} (\Delta_j^2 B - \Delta_j t) \right)^2 = \mathbb{E} \sum_j B_{t_j}^2 (\Delta_j^2 B - \Delta_j t)^2 + 2 \mathbb{E} \sum_j \sum_{k > j} B_{t_j} B_{t_k} (\Delta_j^2 B - \Delta_j t)(\Delta_k^2 B - \Delta_k t) \ .$$

In the first term, independence simplifies calculation, yielding $\mathbb{E} B_{t_j}^2 \, \mathbb{E} (\Delta_j^2 B - \Delta_j t)^2 = O(tn^{-2})$ term-wise, leveraging the variance of a standard χ^2 distribution, which still vanishes in an n-sized sum. The second term can be decomposed

$$\mathbb{E} \left[B_{t_j} B_{t_k} (\Delta_j^2 B - \Delta_j t)(\Delta_k^2 B - \Delta_k t) \right] = \mathbb{E} \left[B_{t_j} B_{t_k} (\Delta_j^2 B - \Delta_j t) \right] \mathbb{E} \left[(\Delta_k^2 B - \Delta_k t) \right] \ ,$$

where the last term vanishes. Of course, $\Delta_j t = n^{-1}$ for all j above.

4

4.3

Let X_t be a stochastic process.

4.3.1

If X_t is a martingale with respect to any filtration $\mathcal{G}_t$, it must be a martingale with respect to its own natural filtration $\mathcal{F}_t$. Measurability and integrability conditions hold immediately, so it suffice to check the conditional martingale property. In addition, by virtue of $X_t \in \mathcal{G}_t$ we must have $\mathcal{F}_t \subset \mathcal{G}_t$ through an inductive argument. Taking $s \le t$, by tower, $\mathbb{E}\left[X_t \mid \mathcal{F}_s\right] = \mathbb{E}\left[\mathbb{E}\left[X_t \mid \mathcal{G}_s\right] \mid \mathcal{F}_s\right] = \mathbb{E}\left[X_s \mid \mathcal{F}_s\right] = X_s$.

4.3.2

By the above, for any martingale X_t, we have for all $t \ge 0$,

$$\mathbb{E}\, X_t = \mathbb{E}\,\mathbb{E}\left[X_t \mid \mathcal{F}_0\right] = \mathbb{E}\, X_0 \ .$$

4.3.3

The above may be satisfied for non-martingales; consider $X_t = t\epsilon$ where ϵ is Rademacher-distributed.

4.4

4.4.1

$X_t = B_t + 4t$ is not a martingale because its expectation is strictly increasing with time (see previous problem).

4.4.2

$X_t = B_t^2$ is not a martingale because its expectation is strictly increasing with time.

4.4.3

$X_t = t^2 B_t - 2 \int_0^t s B_s \, \mathrm{d}s$ is a martingale. Integrability and measurability hold immediately. Then notice

$$\mathbb{E}\left[X_t \mid \mathcal{F}_s\right] = t^2 B_s - 2\int_0^s B_r \, \mathrm{d}r - 2\int_s^t r\, \mathbb{E}\left[B_r \mid \mathcal{F}_s\right] \mathrm{d}r = t^2 B_s - 2\int_0^s B_r \, \mathrm{d}r - 2B_s \int_s^t r \, \mathrm{d}r = X_s \ .$$

4.4.4

$X_t = B_t^{(1)} B_t^{(2)}$ is a martingale, where $B_t^{(i)}$ are independent brownian motions, so the conditional expectation operator will distribute.

5

$M_t = B_t^2 - t$ is a martingale. Measurability and integrability are apparent, and $\mathbb{E}\left[M_t \mid \mathcal{F}_s\right] = B_s^2 + \mathbb{E}\left[2B_s(B_t - B_s) + (B_t - B_s)^2 \mid \mathcal{F}_s\right] - t = M_s$ by splitting up $B_t = B_s + (B_t - B_s)$.

6