

The Itô Formula and the Martingale Representation Theorem

Vladimir Feinberg

September 9, 2021

1 The 1-dimensional Itô Formula

Definition 1.1 (Itô Process). Consider the probability space $(\Omega, \mathcal{H}, \mathbb{P})$ where $\mathcal{H} = \mathcal{H}_\infty$ for a filtration $\{\mathcal{H}_t\}$ for which the Brownian motion B_t is adapted. Let u be an H_t adapted as well, and $L^1([0, T])$ almost surely in its first argument, and $v \in \mathcal{W}_{\mathcal{H}}$, where $\mathcal{W}_{\mathcal{H}}$ is defined in the previous chapter as a relaxation of $\mathcal{V}_{\mathcal{H}}$ to almost surely $L^2([0, T])$ processes for all $T \geq 0$.

Then an Itô process is a process X_t satisfying

$$dX_t = u dt + v dB_t ,$$

which is shorthand for $X_T(\omega) = X_0 + \int_0^T u(t, \omega) dt + \int_0^T v(t, \omega) dB_t$ holding almost surely for all $T \geq 0$.

In certain cases, we can find the functional form of an Itô process through the analogue of the chain rule.

Theorem 1 (The 1-dimensional Itô Formula). *Let X_t be an Itô process satisfying $dX_t = u dt + v dB_t$ and suppose $g \in \mathcal{C}^2(\mathbb{R}_+ \times \mathbb{R})$. Then $Y_t = g(t, X_t)$ is an Itô process satisfying*

$$dY_t = \partial_1 g(t, X_t) dt + \partial_2 g(t, X_t) dX_t + \frac{1}{2} \partial_2^2 g(t, X_t) (dX_t)^2 ,$$

where $(dX_t)^2$ is provocative sugar for dt .

Proof. Letting $V_s = v(s, \omega)$ and $U_s = u(s, \omega)$ and desugaring, we must verify that $g(t, X_t) - g(0, X_0)$ equals

$$I(g, t, X) = \int_0^t \left(\partial_1 g(s, X_s) + U_s \partial_2 g(s, X_s) + \frac{1}{2} V_s^2 \partial_2^2 g(s, X_s) \right) ds + \int_0^t V_s \partial_2 g(s, X_s) dB_s .$$

Before jumping into verification, we must first simplify the problem to work with bounded g and elementary u, v . To argue without loss of generality, we first perform our reductions.

It suffices to consider bounded g . Define the stopping time $\tau_n = \inf \{s > 0 \mid |X_s| \geq n\}$. Define g_n to equal to g , but be supported on only $[0, n] \times [-n, n]$ and $\partial_1 g, \partial_2^2 g$; by continuity and compact support each g_n and its derivatives are bounded.

Defining $\int_0^{t \wedge \tau_n} \dots$ to be $\int_0^t 1 \{s \leq \tau_n\} \dots$, it's clear that $I(g_n, t, X) = I(g, t \wedge \tau_n, X)$. For fixed t , as X_t is $L^2(\mathbb{P} \times [0, t])$ -integrable, $\mathbb{P} \{\tau_n > t\} \rightarrow 1$, which implies that $I(g_n, t, X) = I(g, t \wedge \tau_n, X) \xrightarrow{\text{as}} I(g, t, X)$. Since $g_n \xrightarrow{\text{as}} g$ as well, we can lift a proof of the Itô formula from bounded g to unbounded.

It suffices to consider elementary u, v . Furthermore, we may assume u, v are elementary functions; by integrability and adaptedness (and hence by Monotone Class Theorem) for u and Itô's construction elementary u_n, v_n exist with $dX_t^{(n)} = u_n dt + v_n dB_t$ for which $X_t^{(n)} \xrightarrow{L^1(\mathbb{P})} X_t$, implying almost sure convergence along some subsequence. Then by continuous mapping theorem we have $g(t, X_t^{(n)}) \xrightarrow{\text{as}} g(t, X_t)$. By boundedness of g each integral integrand of $I(g, t, X^{(n)})$ is itself bounded, so by BCT $I(g, t, X^{(n)}) \xrightarrow{L^1(\mathbb{P})} I(g, t, X)$, and therefore through $I(g, t, X^{(n)}) = g(t, X_t^{(n)})$ we have almost sure equality of the limits.

With these amenities out of the way, by application of Taylor's theorem at each $g_j = g(t_j, X_{t_j})$, the difference $g(t, X_t) - g(0, X_0)$ is given by

$$\sum_j \Delta_j g = \sum_j \partial_1 g_j \Delta_j t + \partial_2 g_j \Delta_j X + \frac{1}{2} \partial_1^2 g_j (\Delta_j t)^2 + \frac{1}{2} \partial_2^2 g_j (\Delta_j X)^2 + \partial_1 \partial_2 g_j \Delta_j t \Delta_j X + R_j ,$$

with the remainder $R_j = o\left((\Delta_j t)^2 + (\Delta_j X)^2\right)$.

Under the limit of shrinking partitions, $\sum_j \partial_1 g_j \Delta_j t \rightarrow \int_0^t \partial_1 g(s, X_s) ds$, holding almost surely and then by a limit interchange (relying on g 's boundedness) in $L^1(\mathbb{P})$ as well.

Similarly, we have $\sum_j \partial_2 g_j \Delta_j X \rightarrow \int_0^t \partial_2 g(s, X_s) dB_s$ in $L^2(\mathbb{P})$, since $g \in \mathcal{V}(0, t)$. Next, we leverage the fact that u, v are elementary such that for sufficiently fine grids, there exist constants u_j, v_j such that

$$(\Delta_j X)^2 = u_j^2 (\Delta_j t)^2 + 2u_j v_j \Delta_j t \Delta_j X + v_j^2 (\Delta_j B)^2 .$$

In the original sum, for some bounded c_j , we find that terms $\sum_j c_j \Delta_j t \Delta_j X$, $\sum_j c_j (\Delta_j t)^2$ vanish in $L^2(\mathbb{P})$, the former's second moment being $O(n^{-3})$ and the latter's being $O(n^{-4})$.

This leaves the conspicuous term $\frac{1}{2} \sum_j \partial_2^2 g_j (\Delta_j B)^2$, which we may show tends to $\int_0^t \partial_2^2 g(s, X_s) V_s^2 ds$ in $L^2(\mathbb{P})$ by triangle inequality and comparing to the analogous sequence with $\Delta_j t$ instead. A calculation involving normal moments and leveraging independent increments finds that indeed

$$\mathbb{E} \left[\left(\frac{1}{2} \sum_j \partial_2^2 g_j \left((\Delta_j B)^2 - \Delta_j t \right) \right)^2 \right] \rightarrow 0 .$$

Since every term converges to the desired parts of our integral sum in $L^2(\mathbb{P})$ for some sequence, certainly we can choose a subsequence that converges almost surely. Since the discrete differences always equal $g(t, X_t) - g(0, X_0) = \sum_j \Delta_j g$, the Itô formula equality holds in the limit. \square

Theorem 2 (Stochastic Integration by Parts). *If $f(s, \omega) = f(s)$ is deterministic and of bounded variation in $[0, t]$ such that the [Riemann-Stieltjes integral](#) is well-defined, then*

$$\int_0^t f(s) dB_s = f(t) B_t - \int_0^t B_s df_s$$

Proof. In the case of $f \in \mathcal{C}^2$, this follows from the Itô formula. But for the more general class, a direct proof is needed. For a sequence of shrinking partitions of $[0, t]$,

$$f(t) B_t = \sum_j \Delta_j (fB) = \sum_j B_{t_{j+1}} \Delta_j f + \sum_j f_{t_j} \Delta_j B \xrightarrow{\text{as}} \int_0^t B_s df_s + \int_0^t f(s) dB_s ,$$

which is apparent after adding and subtracting the term $f_{t_j} B_{t_{j+1}}$ within $\Delta_j (fB) = f_{t_{j+1}} B_{t_{j+1}} - f_{t_j} B_{t_j}$. \square

2 The Multi-dimensional Itô Formula

For vector-valued $u : (\mathbb{R}_+, \mathbb{R}^n) \rightarrow \mathbb{R}^n$ and matrix-valued $v : (\mathbb{R}_+, \mathbb{R}^n) \rightarrow \mathbb{R}^{m \times n}$, such that $dX_t = u dt + v dB_t^{(m)}$ has well-defined Itô process components, for a conformal $g \in \mathcal{C}^2$, an analogous multi-dimensional Itô formula may be derived in the same manner as above via Taylor's theorem.

3 The Martingale Representation Theorem

Note that this theorem has a similar multi-dimensional extension.

Theorem 3 (Martingale Representation Theorem). *Suppose M_t is a square-integrable martingale with respect to the filtration generated by Brownian motion. Then there exists a stochastic process $g(s, \omega) \in \mathcal{V}(0, t)$ such that*

$$M_t = \mathbb{E}[M_0] + \int_0^t g(s, \cdot) dB_s ,$$

holding almost surely for all $t \geq 0$.

Proof. The proof proceeds in several steps. First; for any $T > 0$ the set random variables defined by evaluating analytic functions with compact support on R^n on any finite tuple of Brownian motions $(B_{t_1}, \dots, B_{t_n})$ is dense in $L^2(\mathcal{F}_T, \mathbb{P})$, the set of real-valued square-integrable rvs measurable with respect to \mathcal{F}_T .

Second; the linear span of random variables of type $\exp\left(\int_0^T h(t) dB_t - \frac{1}{2} \int_0^T h^2(t) dt\right)$ for all (deterministic) $h \in L^2([0, T])$ is also dense in $L^2(\mathcal{F}_T, \mathbb{P})$ by virtue of spanning the set described above over analytic functions with compact support.

By the density of the class described above, it suffices to consider only random variables of the aforementioned form. For any such Y_T defined by some h , as the final representation equality holds under linear transformation. Since $Y_t = \exp(X_t)$, where $dX_t = -\frac{1}{2}h^2(s) ds + h(s) dB_s$. Applying Itô's formula yields an equality exactly of the desired form $Y_T = 1 + \int_0^T g(s, \cdot) dB_s$. In turn, for any square-integrable, \mathcal{F}_T -measurable random variable, such as M_T , by linearity we may write it as $\mathbb{E}[M_T] + \int_0^T g_T(s, \cdot) dB_s$.

This holds for all T , and $\mathbb{E}[M_T] = \mathbb{E}[M_0]$ uniformly, but we must show that the g_T values are the same across T values. Luckily, this holds due to Exercise 15 of Chapter 3. \square

4 Exercises

4.1

Exercise 4.9 is completed inline in the proof of Itô's formula provided above.