17. ITÔ FORMULA AND PRODUCT RULE FOR SEMIMARTINGALES

With the stochastic integral generalized to integrators that are continuous local martingales, we now revisit concepts introduced earlier in the more specific context. Since the proofs are only simple modifications of the previous ones, we omit most of the details.

17.1 Continuous semimartingales.

We start by generalizing the concept from Definition 13.1:

Definition 17.1 *A stochastic process* $\{X_t\colon t \geqslant 0\}$ *is a* continuous semimartingale *(w.r.t. a given filtration) if there exists a continuous local martingale* $\{M_t\colon t \geqslant 0\}$ *and a continuous adapted process* $\{A_t\colon t \geqslant 0\}$ *with*

$$A_0 = 0 \quad \wedge \quad \forall t \geqslant 0\colon\ V_t^{(1)}(A) < \infty \tag{17.1}$$

such that

$$\forall t \geqslant 0\colon \quad X_t = A_t + M_t \tag{17.2}$$

Note that Lemma 14.3 implies that the processes A and M in the decomposition (17.2) are unique modulo indistinguishability. An example of a continuous semimartingale is a diffusion, in which case

$$A_t = \int_0^t U_s \mathrm{d}s \quad \text{and} \quad M_t = \int_0^t Y_s \mathrm{d}B_s \tag{17.3}$$

for suitable processes U and Y. However, as not all bounded variation functions are absolutely continuous and thus also not all continuous local martingales are stochastic integrals with respect to standard Brownian motion, one can easily generate continuous semimartingales that are not of this form.

Extending the concepts introduced in Definition 13.3, we then get:

Definition 17.2 *Given a continuous semimartingale* X *with decomposition* (17.2)*, for any adapted process* Y *such that, for some* $t \geqslant 0$*,*

$$\int_0^t |Y_s||\mathrm{d}A|_s < \infty \quad \wedge \quad \int_0^t Y_s^2 \mathrm{d}\langle M\rangle_s < \infty \quad \text{a.s.} \tag{17.4}$$

we set

$$\int_0^t Y_s \mathrm{d}X_s := \int_0^t Y_s \mathrm{d}A_s + \int_0^s Y_s \mathrm{d}M_s \tag{17.5}$$

provided Y *has the kind of measurability required for the stochastic integral to exist.*

In order to interpret the above integrals correctly, recall that if $g\colon \mathbb{R}_+ \to \mathbb{R}$ is right-continuous and of bounded variation with $g(0) = 0$, there exists a unique locally-finite signed Borel measure μ_g on $\mathbb{R}_+$ such that $\mu_g([0,t]) = g(t) - g(0)$ for each $t \geqslant 0$. We then write

$$\int_0^t f \mathrm{d}g := \int_{[0,t]} f \mathrm{d}\mu_g \quad \wedge \quad \int_0^t f|\mathrm{d}g| := \int_{[0,t]} f \mathrm{d}|\mu_g| \tag{17.6}$$

where $|\mu_g| := \mu_g^+ + \mu_g^-$ for μ_g^+, μ_g^- denoting the unique positive Borel measures for which the Jordan decomposition $\mu_g = \mu_g^+ - \mu_g^+$ holds with μ_g^+ and μ_g^- mutually singular. Since $|\mu_g|([0,t]) = V_t^{(1)}(g)$, another way to write the above is as

$$\int_0^t f|\mathrm{d}g| = \int_0^t f \mathrm{d}V^{(1)}(g) \tag{17.7}$$

These considerations underlie the otherwise elementary proof of Lemma 17.6 below.

Note that uniqueness of the decomposition (17.2) implies uniqueness of the representation (23.15). In particular, assuming a continuous version of the stochastic integral, the process defined in (17.5) is again a continuous semimartingale (albeit still not necessarily a diffusion, in spite of its integral form).

To simplify future notation, we write

$$\mathscr{S}_{\text{cont}} := \{X\colon \text{ continuous semimartingale}\} \tag{17.8}$$

and, for $X \in \mathscr{S}_{\text{cont}}$ with the decomposition (17.2), put

$$\langle X \rangle_t := \langle M \rangle_t \tag{17.9}$$

We note that, while $\langle X \rangle$ does not have the interpretation as the process that would turn X^2 into a continuous local martingale, it retains the feature that

$$\forall t \geqslant 0\colon \quad V_t^{(2)}(X, \Pi) \underset{\|\Pi\| \to 0}{\xrightarrow{L^2}} \langle X \rangle_t \tag{17.10}$$

For this reason we keep referring to $\langle X \rangle$ as the quadratic variation process associated with X, just as we do for continuous local martingales.

17.2 Integral formulas.

With the above notation in place, we are ready to state:

Theorem 17.3 (Itô formula) *For all $X \in \mathscr{S}_{\text{cont}}$, all $f \in C^2(\mathbb{R})$ and all $t \geqslant 0$,*

$$f(X_t) = f(X_0) + \int_0^t f'(X_s) \mathrm{d}X_s + \frac{1}{2} \int_0^t f''(X_s) \mathrm{d}\langle X \rangle_s \tag{17.11}$$

where all integrals exists thanks to continuity of the integrand.

The proof of this formula is exactly the same as for diffusions (and is in fact likely notationally easier) so we omit it. The second integral on the right is referred to as the *Itô correction*. The expression can be obtained formally by invoking the rule

$$(\mathrm{d}X_t)^2 = \mathrm{d}\langle X \rangle_t \tag{17.12}$$

in the Taylor decomposition $f(X_t + \mathrm{d}X_t) = f(X_t) + f'(X_t)\mathrm{d}X_t + \frac{1}{2} f''(X_t)(\mathrm{d}X_t)^2 + \dots$. Inside the proof this actually shows us as the fact that $(\mathrm{d}M_t)^2 - \mathrm{d}\langle M \rangle_t$ is an infinitesimal increment of a local martingale.

In order to discuss the multivariate version of the Itô formula, as well as other peculiarities of integrals with respect to continuous semimartingales, we need:

Definition 17.4 *Given $M, N \in \mathscr{M}_{\text{loc}}^{\text{cont}}$, denote by $\langle M, N\rangle$ the stochastic process*

$$\langle M, N\rangle_t := \frac{1}{4}\langle M+N\rangle_t - \frac{1}{4}\langle M-N\rangle_t \tag{17.13}$$

We call $\langle M, N\rangle$ the covariation *or* cross-variation *of M and N.*

Clearly, $\langle M\rangle_t = \langle M, M\rangle_t$ so the covariation is just a bilinear form associated with the quadratic form $M \mapsto \langle M\rangle$ via the polarization identity (17.13). We thus readily get:

Lemma 17.5 *Let $M, N \in \mathscr{M}_{\text{loc}}^{\text{cont}}$. Then $\langle M, N\rangle$ is a continuous, adapted process of bounded variation and $\langle M, N\rangle_0 = 0$ such that*

$$\bigl\{M_tN_t - \langle M, N\rangle_t \colon t \geqslant 0\bigr\} \in \mathscr{M}_{\text{loc}}^{\text{cont}}. \tag{17.14}$$

This process is unique up to indistinguishability.

Proof. For existence write

$$M_tN_t = \frac{1}{4}(M_t+N_t)^2 + \frac{1}{4}(M_t-N_t)^2 \tag{17.15}$$

and apply Theorem 14.2 to each of the summands on the right-hand side. The uniqueness then follows from Lemma 14.3. □

For square integrable martingales M and N and $t \geqslant s$ we have $E(M_tN_t|\mathcal{F}_s) = M_sN_s + E((M_t-M_s)(N_t-N_s)|\mathcal{F}_s)$ and so the process $\langle M, N\rangle$ vanishes everywhere if and only if the increments of M and N (over the same intervals) are uncorrelated conditional on the past. In particular, $\langle M, N\rangle$ vanishes if M and N are independent processes with independent increments (such as two independent Brownian motions). The connection to the second variation yields the following pointwise estimates whose proof we leave to a homework exercise:

Lemma 17.6 (Kunita-Watanabe inequality) *For all $M, N \in \mathscr{M}_{\text{loc}}^{\text{cont}}$ we have*

$$\forall t \geqslant 0\colon \quad V_t^{(1)}\bigl(\langle M, N\rangle\bigr) \leqslant \sqrt{\langle M\rangle_t\langle N\rangle_t} \tag{17.16}$$

Moreover, for all $Y, \widetilde{Y} \in \mathcal{V}_M^{\text{loc}} \cap \mathcal{V}_N^{\text{loc}}$,

$$\int_0^t |Y_s\widetilde{Y}_s| \mathrm{d}V_t^{(1)}\bigl(\langle M, N\rangle\bigr) \leqslant \left(\int_0^t Y_s^2 \mathrm{d}\langle M\rangle_s\right)^{1/2}\left(\int_0^t \widetilde{Y}_s^2 \mathrm{d}\langle \widetilde{M}\rangle_s\right)^{1/2} \tag{17.17}$$

holds simultaneously for all $t \geqslant 0$, a.s. (Adaptedness of $Y, \widetilde{Y}$ is not needed.)

If $X = A + M$ and $\widetilde{X} = \widetilde{A} + \widetilde{M}$ are continuous semimartingales, we continue to write

$$\langle X, \widetilde{X}\rangle_t := \langle M, \widetilde{M}\rangle_t \tag{17.18}$$

which is again consistent with the notation (17.9). This becomes useful in:

Lemma 17.7 (Product rule/integration by parts) *Let $X, \widetilde{X} \in \mathscr{S}_{\text{cont}}$. Then for all $t \geqslant 0$,*

$$\int_0^t \widetilde{X}_s \mathrm{d}X_s = \widetilde{X}_tX_t - \widetilde{X}_0X_0 - \int_0^t X_s \mathrm{d}\widetilde{X}_s - \langle \widetilde{X}, X\rangle_t \quad \text{a.s.} \tag{17.19}$$

where the integrals exists thanks to continuity.

Proof. Decomposing $X_t = A_t + M_t$ and $\widetilde{X}_t = \widetilde{A}_t + \widetilde{M}_t$, we now easily check that the formula holds (without the Itô correction) whenever at least one of the processes if of bounded variation. The interesting part thus comes from the case of two local martingales. Here we use that

$$M_t\widetilde{M}_t = \frac{1}{4}(M_t + \widetilde{M}_t)^2 + \frac{1}{4}(M_t - \widetilde{M}_t)^2 \tag{17.20}$$

and note that, if $N \in \mathscr{M}_{\text{loc}}^{\text{cont}}$, the Itô formula gives

$$\mathrm{d}N_t^2 = 2N_t\mathrm{d}N_t + \frac{1}{2}2\,\mathrm{d}\langle N\rangle_t \tag{17.21}$$

Using this for $N_t := M_t \pm \widetilde{M}_t$, polarization identity (17.20) then turns this into

$$\mathrm{d}(M_t\widetilde{M}_t) = 2M_t\mathrm{d}\widetilde{M}_t + 2\widetilde{M}_t\mathrm{d}M_t + \mathrm{d}\langle M, \widetilde{M}\rangle_t \tag{17.22}$$

Integrating we then obtain the desired form. □

17.3 The Fisk-Stratonovich integral.

As observed above, the Itô correction appears in both the analogues of the chain rule (known, under the integral sign, as the Fundamental Theorem of Calculus) as the product rule (known, under the integral sign, as integration by parts). The correction formally arrises from the infinitesimal rule

$$\mathrm{d}X_t\mathrm{d}\widetilde{X}_t = \mathrm{d}\langle X, \widetilde{X}\rangle_t \tag{17.23}$$

which actually encodes a triplet of expressions

$$\mathrm{d}A_t\mathrm{d}\widetilde{A}_t = 0, \quad \mathrm{d}A_t\mathrm{d}M_t = 0 \quad \text{and} \quad \mathrm{d}M_t\mathrm{d}\widetilde{M}_t = \mathrm{d}\langle M, \widetilde{M}\rangle_t \tag{17.24}$$

where $M, \widetilde{M}$ are continuous local martingales and $A, \widetilde{A}$ are processes with bounded variation. There is actually a trick that allows us to somehow absorb the Itô correction into a re-definition of the integral. This is based on:

Definition 17.8 (Fisk-Stratonovich integral) *Given $X, \widetilde{X} \in \mathscr{S}_{\text{cont}}$, for each $t \geqslant 0$ we set*

$$\int_0^t \widetilde{X}_s \circ \mathrm{d}X_s := \int_0^t \widetilde{X}_s\,\mathrm{d}X_s + \frac{1}{2}\langle \widetilde{X}, X\rangle_t \tag{17.25}$$

where the integral on the right is that defined in (17.5).

Indeed, we then have:

Lemma 17.9 *Let $X \in \mathscr{S}_{\text{cont}}$. For each $f \in C^1(\mathbb{R})$ and all $t \geqslant 0$ we then have*

$$\langle f \circ X, X\rangle_t = \int_0^t f'(X_s) \circ \mathrm{d}X_s \quad \text{a.s.} \tag{17.26}$$

In particular, for all $f \in C^2(\mathbb{R})$ and all $t \geqslant 0$ we then have

$$f(X_t) = f(X_0) + \int_0^t f'(X_s) \circ \mathrm{d}X_s \quad \text{a.s.} \tag{17.27}$$

and, for all $\widetilde{X} \in \mathscr{S}_{\text{cont}}$ and all $t \geqslant 0$, also

$$\int_0^t \widetilde{X}_s \circ \mathrm{d}X_s = \widetilde{X}_t X_t - \widetilde{X}_0 X_0 - \int_0^t X_s \circ \mathrm{d}\widetilde{X}_s \tag{17.28}$$

In short, both the chain rule and the product rule hold for the Fisk-Stratonovich integral.

Proof. Assume the decomposition $X_t = A_t + M_t$ from (17.2). By the Itô formula, the martingale part of $(f \circ X)_t$ is the integral $\int_0^t f'(X_s)\mathrm{d}M_s$. Hence,

$$\langle f \circ X, X \rangle_t = \left\langle \int_0^{\cdot} f'(X_s)\mathrm{d}M_s, M \right\rangle_t \tag{17.29}$$

Using that

$$\int_0^t f'(X_s)\mathrm{d}M_s \pm M_t = \int_0^t \bigl[f'(X_s) \pm 1\bigr]\mathrm{d}M_s \quad \text{a.s.} \tag{17.30}$$

and so, by (16.30),

$$\left\langle \int_0^t f'(X_s)\mathrm{d}M_s \pm M_t \right\rangle_t = \int_0^t \bigl[f'(X_s) \pm 1\bigr]^2 \mathrm{d}\langle M \rangle_s \quad \text{a.s.} \tag{17.31}$$

The polarization identity gives

$$\left\langle \int_0^{\cdot} f'(X_s)\mathrm{d}M_s, M \right\rangle_t = \int_0^t f'(X_s)\mathrm{d}\langle M \rangle_s \quad \text{a.s.} \tag{17.32}$$

In light of (17.9), this is (17.26). The remaining claims then follow by identification of terms in (17.11) and (17.19) via (17.25). □

The integral (17.25) subsumes that introduced in (7.26). As it turns out the Fisk-Stratonovich integral actually arises as the limit of Riemann sums under a version of the midpoint rule. This comes in:

Lemma 17.10 *Let $X, \widetilde{X} \in \mathscr{S}_{\text{cont}}$. Then for any $t \geqslant 0$ and any sequence of partitions $\{\Pi_n\}$ of $[0,t]$, where $\Pi_n = \{0 = t_0 < t_1 < \dots < t_m(n) = t\}$, if $\|\Pi_n\| \to 0$ then*

$$\sum_{i=1}^{m(n)} \frac{\widetilde{X}_{t_i} + \widetilde{X}_{t_{i-1}}}{2}(X_{t_i} - X_{t_{i-1}}) \xrightarrow[n\to\infty]{P} \int_0^t \widetilde{X}_s \circ \mathrm{d}X_t \tag{17.33}$$

Note that, unlike the Itô integral, the use of the midpoint rule mucks up the martingale property of the Riemann sums so it is not surprising that the resulting Fisk-Stratonovich integral is not a martingale either (that is, unless the covariation of the two processes vanishes). The Fisk-Stratonovich integral is not good for theory development, where Itô integral is of prime importance, but it is often more useful in complex calculations where having the "usual" calculus rules is more efficient. We leave the proof of Lemma 17.10 to a homework exercise.

17.4 Multivariate Itô formula.

As our final topic in this section, we will generalize the Itô formula to functions of time and multiple continuous semimartingales. This subsumes Theorem 13.8 where this was stated for diffusions.

Theorem 17.11 (Multivariate Itô formula) *Let* $X^{(1)}, \dots, X^{(d)} \in \mathscr{S}_{\text{cont}}$ *and suppose that* $f\colon \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}$ *is of type* $C^1 \times C^2$*. Denoting* $X_t := (X^{(1)}_t, \dots, X^{(d)}_t)$*, for all* $t \geqslant 0$ *we have*

$$\begin{aligned} f(t, X_t) = f(0, X_0) &+ \int_0^t \frac{\partial f}{\partial s}(s, X_s)\mathrm{d}s \\ &+ \sum_{i=1}^m \int_0^t \frac{\partial f}{\partial x_i}(s, X_s)\mathrm{d}X^{(i)}_s + \frac{1}{2}\sum_{i,j=1}^d \int_0^t \frac{\partial^2 f}{\partial x_i \partial x_j}(s, X_s)\mathrm{d}\langle X^{(i)}, X^{(j)}\rangle_s \quad \text{a.s.} \end{aligned} \tag{17.34}$$

where the integrals exists by the assumed continuity of the derivatives.

We will not give the proof of this as it follows very much the same lines at that of its univariate version. The result can formally be obtained by invoking the rules

$$(\mathrm{d}t)^2 = 0, \quad \mathrm{d}t\mathrm{d}X^{(i)}_t = 0 \quad \text{and} \quad \mathrm{d}X^{(i)}_t \mathrm{d}X^{(j)}_t = \mathrm{d}\langle X^{(i)}, X^{(j)}\rangle_t \tag{17.35}$$

in the expansion of $f(t + \mathrm{d}t, X_t + \mathrm{d}X_t)$ into a Taylor polynomial. The reader should notice the efficiency of the semimartingale notation compared to that in Theorem 13.8.

Further reading: Section 3.3A of Karatzas-Shreve