B8.2 Continuous Martingales and Stochastic Calculus

Itô's formula and its applications

Itô's formula and integration by parts

Samuel Cohen
Hilary Term 2021

UNIVERSITY OF OXFORD
Mathematical Institute

Oxford Mathematics

We already saw that the stochastic integral of Brownian motion with respect to itself did not behave as we would expect from Newtonian calculus.

So what are the analogues of integration by parts and the chain rule for stochastic integrals?

Proposition (Integration by parts)

If X and Y are two continuous semimartingales then

$$\begin{aligned} X_t Y_t &= X_0 Y_0 + \int_0^t X_s \mathrm{d} Y_s + \int_0^t Y_s \mathrm{d} X_s + \langle X, Y \rangle_t, \quad t \geq 0 \quad a.s. \\ &= X_0 Y_0 + (X \bullet Y)_t + (Y \bullet X)_t + \langle X, Y \rangle_t. \end{aligned}$$

Remark

Comparing with the finite variation case with jumps (Proposition 7.8), we see that the 'product of jumps' term has become the 'quadratic variation' term.

Proof.

Fix t and let π^n be a sequence of partitions of $[0,t]$ with mesh converging to zero. Note that

$$X_t Y_t - X_s Y_s = X_s(Y_t - Y_s) + Y_s(X_t - X_s) + (X_t - X_s)(Y_t - Y_s)$$

so for any n

$$\begin{aligned} X_t Y_t - X_0 Y_0 = \sum_{t_i \in \pi^n} \Big(& X_{t_i}(Y_{t_{i+1}} - Y_{t_i}) + Y_{t_i}(X_{t_{i+1}} - X_{t_i}) \\ & + (X_{t_{i+1}} - X_{t_i})(Y_{t_{i+1}} - Y_{t_i}) \Big) \\ \longrightarrow (X \bullet Y)_t + (Y \bullet X)_t + \langle X, Y \rangle_t & \quad \text{as } n \to \infty. \end{aligned}$$

$\square$

Theorem (Itô's formula)

Let $X^1, \ldots, X^d$ be continuous semimartingales and $F : \mathbb{R}^d \to \mathbb{R}$ a C^2 function. Then $(F(X_t^1, \ldots, X_t^d) : t \geq 0)$ is a continuous semimartingale and up to indistinguishability

$$
\begin{aligned}
F(X_t^1, \ldots, X_t^d) = & F(X_0^1, \ldots, X_0^d) + \sum_{i=1}^{d} \int_0^t \frac{\partial F}{\partial x^i}(X_s^1, \ldots, X_s^d) \mathrm{d}X_s^i \\
& + \frac{1}{2} \sum_{1 \leq i,j \leq d} \int_0^t \frac{\partial^2 F}{\partial x^i \partial x^j}(X_s^1, \ldots, X_s^d) \mathrm{d}\langle X^i, X^j \rangle_s.
\end{aligned} \tag{1}
$$

In particular, for $d = 1$, we have

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

Proof

Let $X^i = X_0^i + M^i + A^i$ be the semimartingale decomposition of X^i and denote by V^i the total variation process of A^i. Let

$$\tau_r^i = \inf\{t \geq 0 : |X_t^i| + V_t^i + \langle M^i \rangle_t > r\},$$

and $\tau_r = \min\{\tau_r^i, i = 1, \ldots, d\}$. Then $(\tau_r)_{r \geq 0}$ is a family of stopping times with $\tau_r \uparrow \infty$. It is sufficient to prove (1) up to time τ_r.

We will prove that the result holds for polynomials and then the full result follows by approximating C^2 functions by polynomials.

First note that it is obvious that the set of functions for which the formula holds is a vector space containing the functions $F \equiv 1$ and $F(x_1, \ldots, x_d) = x_i$ for $i \leq d$.

We now check that if (1) holds for two functions F and G, then it holds for the product FG. Integration by parts yields

$$F_t G_t - F_0 G_0 = \int_0^t F_s \mathrm{d} G_s + \int_0^t G_s \mathrm{d} F_s + \langle F, G \rangle_t. \tag{2}$$

By associativity of stochastic integration, and because (1) holds for G,

$$\int_0^t F_s \mathrm{d} G_s = \sum_{i=1}^d \int_0^t F(X_s) \frac{\partial G_s}{\partial x^i} \mathrm{d} X_s^i + \frac{1}{2} \sum_{1 \le i,j \le d} \int_0^t F(X_s) \frac{\partial^2 G_s}{\partial x^i \partial x^j} \mathrm{d} \langle X^i, X^j \rangle_s,$$

with a similar expression for $\int_0^t G_s \mathrm{d} F_s$.

Using the fact that (1) holds for F and G, we also have

$$\langle F, G\rangle_t = \sum_{i=1}^{d}\sum_{j=1}^{d}\int_0^t \frac{\partial F_s}{\partial x^i}\frac{\partial G_s}{\partial x^j}\mathrm{d}\langle X^i, X^j\rangle_s.$$

Substituting these into (2), we obtain Itô's formula for FG.

To pass to a general C^2 function F, the Stone–Weierstraß theorem (see appendix) allows us to approximate the second derivative of F uniformly on compacts by a polynomial (and hence F' and F are also uniformly approximated on compacts).

Using the dominated convergence theorem (and the fact that everything is nicely bounded up to time τ_r), we have the result up to time τ_r, and then we send $r \to \infty$. $\square$