Statistics 212: Lecture 17 (April 7, 2025)

Ito's lemma

Instructor: Mark Sellke

Scribe: Georgi Ivanov

1 Ito's formula

Proposition 17.1 If $f \in C_b^\infty(\mathbb{R})$ (i.e. all derivatives of f are uniformly bounded on $\mathbb{R}$), then:

$$f(B_t) = f(B_0) + \int_0^t f'(B_s)\,dB_s + \frac{1}{2}\int_0^t f''(B_s)\,ds, \tag{1}$$

or written in the informal differential form:

$$df(B_t) = f'(B_t)\,dB_t + \frac{f''(B_t)}{2}\,dt.$$

(Note: as mentioned below, the proof uses boundedness of f''', but after the fact you can deduce the result for C_b^2 by approximation, i.e. you just need that (f, f', f'') are bounded and continuous. And if f is time-dependent you also need $\partial_t f$ to be bounded and continuous.)

1.1 Proof

The idea is to write the Taylor expansion and show that it is valid. First, let us fix a small mesh partition $\{t_0 = 0 \le t_1 < t_2 < \cdots < t_k = t\}$, with $\epsilon = \max_i |t_{i+1} - t_i|$. We can express the difference as a a telescoping sum of the Brownian increments and apply the second-order Taylor expansion with the Lagrange remainder $\theta_i \in [t_i, t_{i+1}]$:

$$\begin{aligned}
f(B_t) - f(B_0) &= \sum_{i=0}^{k-1} \left(f(B_{t_{i+1}}) - f(B_{t_i}) \right) \\
&= \sum_{i=0}^{k-1} f'(B_{t_i})(B_{t_{i+1}} - B_{t_i}) && \text{(T1)} \\
&\quad + \sum_{i=0}^{k-1} \frac{1}{2} f''(B_{t_i})(B_{t_{i+1}} - B_{t_i})^2 && \text{(T2)} \\
&\quad + \sum_{i=0}^{k-1} \frac{1}{6} f'''(\theta_i)(B_{t_{i+1}} - B_{t_i})^3 && \text{(T3)}
\end{aligned}$$

As we take the limit $\epsilon \to 0$, the first term $(T1) \to \int_0^t f'(B_s)\,dB_s$, the second term $(T2) \to \frac{1}{2}\int_0^t f''(B_s)\,ds$, and the last $(T3) \to 0$, where all limits are in L^2. Let us verify this explicitly:

Term 1 ($T1$)

$$T1 = \sum_{i=0}^{k-1} f'(B_t)(B_{t_i+1} - B_{t_i}) = \int_0^t f'(B_{\tilde{s}})dB_{\tilde{s}}, \quad \tilde{s} = \tilde{s}(s) = t_i, \quad \text{if } s \in [t_i, t_{i+1}]$$

The integral term is piecewise constant in time and progressively measurable. We want to show that this is $\approx \int_0^t f'(B_s)dB_s$. By the Ito isometry we can verify that the L^2 error between the two integrants goes to zero:

$$\mathbb{E}\int_0^t \left(f(B_s) - f(B_{\tilde{s}})\right)^2 ds \to 0$$

as $\epsilon \to 0$. The term is is bounded almost surely because $f \in C_b^\infty$ is bounded and uniformly continuous, s is uniformly Lipschitz, and B_t is uniformly continuous.

Term ($T2$)

$$T2 = \sum_i f''(B_{t_i})(B_{t_i+1} - B_{t_i})^2 \to^{L^2} \int_0^t f''(B_s)ds$$

We know $\sum f''(B_{t_i}))(t_{i+1} - t_i) \to \int_0^t f''(B_s)ds$ to the RHS above by the classical Riemann sum form. As such, we want to show that the two sums are close in the L^2 norm. The idea is to use a martingale argument to show concentration. Let $M_0, \cdots M_k$ be defined as:

$$M_j = \sum_{i=0}^{j} f''(B_{t_i}) \times \left((B_{t_{i+1}} - B_{t_i})^2 - (t_{i+1} - t_i)\right),$$

where M_j forms a martingale. We can compute the variance as:

$$\mathbb{E}\left[M_{k-1}^2\right] = \sum_{i=0}^{k-1} \mathbb{E}\left[f''(B_{t_i})^2 \times \left((B_{t_{i+1}} - B_{t_i})^2 - (t_{i+1} - t_i)\right)^2\right]$$

We are going to use the bound $(a-b)^2 \le 2(a^2 + b^2)$, absorbing the 2 in $C(f)$, along with the fact that most of the squared terms are orthogonal. As such, we get:

$$\mathbb{E}\left[M_{k-1}^2\right] \le C(f)\sum_{i=0}^{k-1} \mathbb{E}\left[(B_{t_{i+1}} - B_{t_i})^4 + (t_{i+1} - t_i)^2\right] = C(f)\sum_{i=0}^{k-1} \mathbb{E}\left[(B_{t_{i+1}} - B_{t_i})^4\right] + C(f)\sum_{i=0}^{k-1}(t_{i+1} - t_i)^2$$

First, observe that $\mathbb{E}[(B_{t_{i+1}} - B_{t_i})^4] = 3(t_{i+1} - t_i)^2$. Now, we just need $\sum_{i=0}^{k-1}(t_{i+1} - t_i)^2 \to 0$ as $\epsilon \to 0$. One convinient way to estimate these sums is to use the fact that

$$\max_i |t_{i+1} - t_i| \times \sum_{i=0}^{k-1}(t_{i+1} - t_i)$$

is of order $\epsilon t \to 0$ as $\epsilon \to 0$.

Term T3 ($T3$)

Here, we just need the $\sum \mathbb{E}|B_{t_{i+1}} - B_{t_i}|^3 \to 0$ as $\epsilon \to 0$. Since $\mathbb{E}|B_{t_{i+1}} - B_{t_i}|^3 \le O(|t_{i+1} - t_i|^{3/2})$ we similarly get

$$\sum \mathbb{E}|B_{t_{i+1}} - B_{t_i}|^3 \le C\sum_i |t_{i+1} - t_i|^{3/2} \le C\sqrt{\epsilon}\sum_i |t_{i+1} - t_i| \le Ct\sqrt{\epsilon}.$$

With this, as we take the limit of $(T1) + (T2) + (T3)$, we get (1). $\square$

2 General Ito's formula

Proposition 17.2 More generally, suppose that $dX_t = \sigma_t dB_t + v_t dt$. We want to know what $df(t, X_t)$ is (in stochastic form). We have:

$$df(t, X_t) = \partial_x f(t, X_t)\sigma_t dB_t + \frac{1}{2}\sigma_t^2 \partial_{xx} f(t, X_t) dt + \partial_x f(t, X_t) v_t dt + \partial_t f(t, X_t) dt \tag{2}$$

The proof is equivalent to our earlier approach by using telescoping sums and bounding the L^2 error terms. The usual assumption is $f \in C_b^{1,2}$ (our proof used f''', but given the statement of Ito's formula for smooth functions, you get it for $C_b^{1,2}$ by approximation).

An extension of the formula also works for $f(B_t) = |B_t|$. Here, the keywords are "Tanaka's formula" and "Brownian local time" (p. 209 in the Mörters and Perez). (The short summary is that $\int_0^t f''(B_s) ds$ is the integral of f'' against the "occupation measure" $\mu_{[0,t]}$ of Brownian motion, defined by $\mu_{[0,t]}(A) = \int_0^t 1_{B_s \in A} ds$. One can show $\mu_{[0,t]}$ has a continuous density, so one can integrate a Dirac delta (such as the 2nd derivative of absolute value) against it by a continuity argument.)

For the multidimensional case $\mathbb{R}^d \to \mathbb{R}$, suppose that $f(B_t^1, B_t^2, \cdots, B_t^d)$. Then, the Ito correction is, instead of f'', the Laplacian $\Delta f = \sum_{i=1}^d (\partial_{x_i})^2 f$.

2.1 Examples

(a) $f(t,x) = tx$

$$df(t,x) = \partial_x f dB_t + \partial_t f dt = t dB_s + x dt$$

$$tB_t = \int_0^t s dB_s + \int_0^t B_s ds$$

(b) $f(x) = x^2$

$$df(B_t) = 2B_t dB_t + dt$$

$$B_t^2 = 2\int_0^t B_s dB_s + t \Leftrightarrow B_t^2 - t = 2\int_0^t B_s dB_s$$

(c) Fix $\alpha > 0$, consider a solution to the following Bessel process:

$$dX_t = dB_t + \frac{\alpha}{X_t} dt, \quad X_0 = 1$$

For the last example, we will stop if we reach 0. The question is whether it does hit 0. The idea is to find γ so $f(X_t) = X_t^\gamma$ is a martingale. We want an expression for dX_t^γ, so we will have to use the general formula (2):

$$dX_t^\gamma = \gamma X_t^{\gamma-1} dB_t + \left(\alpha\gamma X_t^{\gamma-2} + \frac{\gamma(\gamma-1)}{2} X_t^{\gamma-2}\right) dt$$

We want the dt term to be 0, which occurs when $\gamma = 1 - 2\alpha$, i.e $\frac{1-\gamma}{2} = \alpha$. $Y_t = F(X_t)$ is a continuous martingale with $Y_0 = 1$ and $Y_t \geq 0, \forall t$. By OST $\mathbb{P}[\max_t Y_t \geq y_\epsilon] \leq \epsilon$. If $\gamma \leq 0$, then $X_t \to 0 \Leftrightarrow Y_t \to \infty$ which doesn't happen because of OST, i.e X_t never reaches 0. This means we need $\alpha > \frac{1}{2}$.

The conclusion that X_t never reaches 0 for $\alpha > 1/2$ is correct, but actually X_t^γ is not necessarily a martingale! What goes wrong is the first term $\gamma X_t^{\gamma-1} dB_t$, which can be unbounded (it will be bounded in time almost surely, but can be arbitrarily large depending on the Brownian path). So the stochastic integral is not even defined in the L^2 sense we considered so far. However it seems intuitively clear that nothing should go wrong if we never actually reach 0.

The resolution is to consider a larger class of integrands which are "locally L^2" and progressively measurable. An integrand $H(t,\omega)$ is locally L^2 if there exists increasing sequence of stopping times $0 \le \tau_1 \le \tau_2 \le \cdots$ which tend to infinity $\lim_{k\to\infty} \tau_k = \infty$ almost surely, such that:

$$\forall k, \quad H(t,\omega)\mathbb{1}_{t\le\tau_k} \in L^2, i.e \quad \mathbb{E}\int_0^{\tau_k} H(t,\omega)^2 dt < \infty.$$

This means we can integrate up to each stopping time, hence we can integrate all the way to infinity. Namely, consider for each k the process

$$t \mapsto \int_0^t H(s,\omega)\mathbb{1}_{s\le\tau_k} ds.$$

Since $\tau_k \to \infty$, the latter provides us a def. of

$$t \mapsto \int_0^t H(s,\omega) dB_s = \lim_{k\to\infty} \int_0^t H(s,\omega)\mathbb{1}_{s\le\tau_k} dB_s \text{ a.s.}$$

(The homework will ask us to check that this construction is well-defined.)

In the example we saw of the Bessel process, we can say τ_k is the hitting time of $\{1/k\}$ (and then X_t and Y_t stop at this time as we set the integrand to 0 afterward). The use of the optional stopping theorem was acceptable because for each k, we had $\mathbb{E}[Y_t^k] = Y_0 = 1$, and so Fatou's lemma gives $\mathbb{E}[Y_t] \le 1$.