Week 6
Ito's lemma for Brownian motion

Jonathan Goodman

October 22, 2012

1 Introduction to the material for the week

sec:intro

Ito's lemma is the big thing this week. It plays the role in stochastic calculus that the fundamental theorem of calculus plays in ordinary calculus. Most actual calculations in stochastic calculus use some form of Ito's lemma. Ito's lemma is one of a family of facts that make up the *Ito calculus*. It is an analogue for stochastic processes of the ordinary calculus of Leibnitz and Netwon. We use it both as a language for expressing models, and as a set of tools for reasoning about models.

For example, suppose N_t is the number of bacteria in a dish (a standard example in beginning calculus). We model N_t in terms of a growth rate, r. In a small increment of time dt, the model is that N increases by an amount $dN_t = rN_t dt$. Calculus allows us to express N_t as $N_t = N_0 e^{rt}$. The "Ito's lemma" of ordinary calculus gives $df(t) = f'(t)dt$. For us, this is $d(N_0 e^{rt} = rN_0 e^{rt} = rN_t$.

Here is a similar example for a stochastic process X_t that could model a stock price. We suppose that in the time interval dt that X_t changes by a random amount whose size is proportional to X_t. In stock terms, the probability to go from 100 to 102 is the same as the probability to go from 10 to 10.2. A simple way to do this is to make dX proportional to X_t and dW_t, as in $dX_t = \sigma X_t dW_t$. The differentials are all forward looking, so $dX_t = X_{t+dt} - X_t$ and $dW_t = W_{t+dt} - W_t$ with $dt > 0$. The Ito lemma for the Ito calculus is (using subscripts for partial derivatives) $d(f(W_t,t)) = f_w(W_t,t)dW_t + \frac{1}{2} f_w(W_t,t)dt + f_t(W_t,t)dt$. The solution is $f(w,t) = x_0 e^{\sigma w - \sigma^2 t/t}$. We check this using $f_w = \sigma f$, $f_{ww} = \frac{\sigma^2}{2} f$, and $f_t = \frac{-\sigma^2}{2} f$. Therefore, if $X_t = f(W_t,t)$, then $dX_t = \sigma X_t dW_t$ as desired.

Ito's lemma for this week is about the time derivative of stochastic processes $f(W_t,t)$, where $f(w,t)$ is a differentiable function of its arguments. The Ito differential is

$$df = f(W_{t+dt}, t+dt) - f(W_t,t) \; .$$

This is the change in f over a small increment of time dt. If you integrate the Ito differential of f, you get the change in f. If X_t is any process, then

$$X_b - X_a = \int_a^b dX_s \; . \qquad (1)$$

eq:di

This is the way to show that something is equal to dX_t, you put your differential on the right, integrate, and see whether you get the left side. In particular, the differential formula $dX_t = \mu_t dt + \sigma_t dW_t$, means that

$$X_b - X_a = \int_a^b \mu_s \, ds + \int_a^b \sigma_s \, dW_s \; . \tag{2}$$

eq:dii

The first integral on the right is an ordinary integral. The second is the Ito integral from last week. The Ito integral is well defined provided σ_t is an adapted process.

Ito's lemma for Brownian motion is

$$df(W_t,t) = \partial_w f(W_t,t)dW_t + \frac{1}{2}\partial_w^2 f(W_t,t)dt + \partial_t f(W_t,t)dt \; . \tag{3}$$

eq:ild

An informal derivation starts by expanding df in Taylor series in dW and dt up to second order in dW and first order in dt,

$$df = \partial_w f \, dW + \frac{1}{2}\partial_w^2 f \, (dW)^2 + \partial_t f \, dt \; .$$

We get (3) from this using $(dW_t)^2 = dt$. The formula $dW_t)^2 = dt$ cannot be exactly true, because $(dW_t)^2$ is random and dt is not random. It is true that $\mathrm{E}\left[\,(dW_t)^2 | \mathcal{F}_t\right] = dt$, but Ito's lemma is about more than expectations.

The real theorem of Ito's lemma, in the spirit of (2), is

$$\begin{aligned} &f(W_b,b) - f(W_a,a) \\ &\qquad = \int_a^b \partial_w f(W_t,t)dW_t + \int_a^b \left(\frac{1}{2}\partial_w^2 f(W_t,t) + \partial_t f(W_t,t)\right) dt \end{aligned} \tag{4}$$

eq:ili

Everything here is has been defined. The second integral on the right is an ordinary Riemann integral. The first integral on the right is the Ito integral defined last week. We give an informal proof of this in Section 2.

You see the convenience of Ito's lemma by re-doing the example from last week

$$X_t = \int_0^t W_s dW_s \; .$$

A first guess from ordinary calculus might be $X_t = \frac{1}{2}W_t^2$. Let us take the Ito differential of $\frac{1}{2}W_t^2$. This is $df(W_t,t)$, where $f(w,t) = \frac{1}{2}w^2$, and $\partial_w f(w,t) = w$, and $\frac{1}{2}\partial_w^2 f(w,t) = \frac{1}{2}$. Therefore, (3) gives

$$d\left(\frac{1}{2}W_t^2\right) = W_t dW_t + \frac{1}{2}dt \; .$$

Therefore,

$$\begin{aligned} \frac{1}{2}W_t^2 - \frac{1}{2}W_0^2 &= \int_0^t W_s dW_s + \frac{1}{2}\int_0^t ds \\ &= \int_0^t W_s dW_s + \frac{1}{2}t \; . \end{aligned}$$

You just rearrange this and recall that $W_0 = 0$, and you get the formula from Week 5:

$$X_t = \int_0^t W_s dW_s = \frac{1}{2}W_t^2 - \frac{1}{2}t \; .$$

This is quicker than the telescoping sum stuff from Week 5.

Ito's lemma gives a convenient way to figure out the backward equation for many problems. Ito's lemma and the martingale (mean zero) property of Ito integrals work together to tell you how to evaluate conditional expectations. Consider the Ito integral

$$X_T = \int_0^T g_s dW_s \; .$$

Then

$$\mathrm{E}[\, X_T \mid \mathcal{F}_t] = \mathrm{E}\left[\int_0^t g_s dW_s \mid \mathcal{F}_t \right] + \mathrm{E}\left[\int_t^T g_s dW_s \mid \mathcal{F}_t \right]$$

The first term is completely known at time t, so the expectation is irrelevant. The second term is zero, because dW_s is in the future of g_s and $\mathcal{F}_t$. Therefore

$$\mathrm{E}\left[\int_0^T g_s dW_s \mid \mathcal{F}_t \right] = \int_0^t g_s dW_s \; .$$

Now suppose $f(w,t)$ is the value function

$$f(w,t) = \mathrm{E}[\, V(W_T) \mid W_t = w] \; .$$

The integral form of Ito's lemma (4)

$$\begin{aligned} V(W_T) - f(W_t,t) &= \int_t^T df(W_s,s) \\ &= \int_t^T \partial_w f(W_s,s) dW_s + \int_t^T \left(\partial_t f(W_s,s) + \frac{1}{2}\partial_w^2 f(W_s,s) \right) ds \end{aligned}$$

Take the conditional expectation in $\mathcal{F}_t$. Looking on the left side, we have

$$\mathrm{E}[\, V(W_T) \mid \mathcal{F}_t] = f(W_t,t) \; ,$$

which is an equivalent definition of the value function. Clearly, $\mathrm{E}[\, f(W_t,t) \mid \mathcal{F}_t] = f(W_t,t)$. Therefore you get zero on the left. The conditional expectation of the Ito integral on the right also vanishes, as we said just above. Therefore

$$\mathrm{E}\left[\int_t^T \left(\partial_t f(W_s,s) + \frac{1}{2}\partial_w^2 f(W_s,s) \right) ds \mid \mathcal{F}_t \right] = 0 \; .$$

The simplest way for this to happen is for the integrand to vanish identically. The equation you get by setting the integrand to zero is

$$\partial_t f + \frac{1}{2}\partial_w^2 f = 0 \; .$$

This is the backward equation we derived in Week 4. The difference here is that you don't have to think about what you're doing here. All the hard thinking (the mathematical analysis) goes into Ito's lemma. Once you are liberated from thinking hard, you can easily derive backward equations for many other situations.

2 Informal proof of Ito's lemma

sec:p

The theorem of Ito's lemma is the integral formula (4). We will prove it under the assumption that $f(w,t)$ is a differentiable function of its arguments up to third derivatives. We assume all mixed partial derivatives up to that order exist and are bounded. That means $\left|\partial_w^3 f(w,t)\right| \leq C$, and $\left|\partial_t^2 f(w,t)\right| \leq C$, and $\left|\partial_w^2 \partial_t f(w,t)\right| \leq C$, and so on.

We use the notation of Week 5, with $\Delta t = 2^{-m}$, and $t_j = j\Delta t$. The change in any quantity from t_j to t_{j+1} is $\Delta(**)_j$. We use the subscript j for t_j, as in W_j instead of W_{t_j}. For example, $\Delta f_j = f(W_j + \Delta W_j, t_j + \Delta t) - f(W_j, t_j)$. In this notation, the left side of (4) is

$$f(W_b, b) - f(W_a, a) \approx \sum_{a \leq t_j < b} \Delta f_j \ . \qquad (5)$$ eq:dfs

The right side is a telescoping sum, which is equal to the left side if $b = n\Delta t$ and $a = m\Delta t$ for some integers $m < n$. When Δt and ΔW are small, there is a Taylor series approximation of Δf_j. The leading order terms in the Taylor series combine to form the integrals on the right of (4). The remainder terms add up to something that goes to zero as $\Delta t \to 0$.

Suppose w and t are some numbers and Δw and Δt are some small changes. Define $\Delta f = f(w + \Delta w, t + \Delta t) - f(w,t)$. The Taylor series, up to the order we need, is

$$\Delta f = \partial_w f(w,t)\Delta w + \tfrac{1}{2}\partial_w^2 f(w,t)\Delta w^2 + \partial_t f(w,t)\Delta t \qquad (6)$$ eq:ts

$$+ O\left(\left|\Delta w^3\right|\right) + O\left(\left|\Delta w\right| \Delta t\right) + O\left(\left|\Delta t^2\right|\right) \ . \qquad (7)$$ eq:et

The big O quantities on the second line refer to things bounded by a multiple of what's in the big O, so $O\left(\left|\Delta w^3\right|\right)$ means: "some quantity Q so that there is a C with $|Q| \leq C\left|\Delta w^3\right|$". The error terms on the second line correspond to the highest order neglected terms in the Taylor series. These are (constants omitted) $\partial_w^3 f(w,t)\Delta w^3$, and $\partial_w \partial_t f(w,t)\Delta w \Delta t$, and $\partial_t^2 f(w,t)\Delta t^2$. The Taylor remainder theorem tells us that if the derivatives of the appropriate order are bounded (third derivatives in this case), then the errors are on the order of the neglected terms.

The sum on the right of (5) now breaks up into six sums, one for each term on the right of (6) and (7):

$$\sum_{a \leq t_j < b} \Delta f_j = S_1 + S_2 + S_3 + S_4 + S_5 + S_6 \ .$$

We consider them one by one. It does not take long.

The first is

$$S_1 = \sum_{a \le t_j < b} \partial_w f(W_j, t_j) \Delta W_j \ .$$

In the limit $\Delta t \to 0$ (more precisely, $m \to \infty$ with $\Delta t = 2^{-m}$), this converges to

$$\int_a^b \partial_w f(W_s, s) dW_s \ .$$

The second is

$$S_2 = \sum_{a \le t_j < b} \tfrac{1}{2} \partial_w^2 f(W_j, t_j) \Delta W_j^2 \ . \tag{8}$$

eq:is

This is the term in the Ito calculus that has no analogue in ordinary calculus. We come back to it after the others. The third is

$$S_3 = \sum_{a \le t_j < b} \partial_t f(W_j, t_j) \Delta t \ .$$

As $\Delta t \to 0$ this one converges to

$$\int_a^b \partial_t f(W_s, s) \, ds \ .$$

The first error sum is

$$|S_4| \le C \sum_{a \le t_j < b} \left| \Delta W_j^3 \right| \ .$$

This is random, so we evaluate its expected value. We know from experience that $\mathrm{E}\big[\left| \Delta W_j^3 \right| \big]$ scales like $\Delta t^{3/2}$, which is one half power of Δt for each power of ΔW. Therefore

$$\mathrm{E}[\, S_4] \le C \sum_{a \le t_j < b} \Delta t^{3/2} = C \Delta t^{1/2} \sum_{a \le t_j < b} \Delta t = C(b-a) \Delta t^{1/2} \ .$$

The second error term goes the same way, as $\mathrm{E}[\, |\Delta W_j| \, \Delta t]$ also scales as $\Delta t^{3/2}$. The last error term has

$$|S_6| \le C \sum_{a \le t_j < b} \Delta t^2 = C(b-a) \Delta t \ .$$

It comes now to the sum (eq:is 8). The $\left(\Delta W_j^2 \right) \leftrightarrow \Delta t$ connection suggests we write

$$\left(\Delta W_j \right)^2 = \Delta t + R_j \ ,$$

Clearly

$$\mathrm{E}[\, R_j \mid \mathcal{F}_j] = 0 \ \ , \quad \text{and } \ \mathrm{E}\left[R_j^2 \mid \mathcal{F}_j \right] = \mathrm{var}(R_j \mid \mathcal{F}_j) = 2 \Delta t^2 \ .$$

Now,

$$
\begin{aligned}
S_2 = & \sum_{a \le t_j < b} \tfrac{1}{2}\partial_w^2 f(W_j, t_j)\Delta t \; + \sum_{a \le t_j < b} \tfrac{1}{2}\partial_w^2 f(W_j, t_j) R_j \\
= & \qquad S_{2,1} \qquad\qquad + \qquad S_{2,2} \; .
\end{aligned}
$$

The first term converges to the Riemann integral

$$
\int_a^b \tfrac{1}{2}\partial_w^2 f(Ws, s)\, ds \; .
$$

The second term converges to zero almost surely. We see this using the now familiar trick of calculating $\mathrm{E}\big[\, S_{2,2}^2 \big]$. This becomes a double sum over t_j and t_k. The off diagonal terms, the ones with $j \neq k$ vanish. If $j > k$, we see this as usual:

$$
\begin{aligned}
& \mathrm{E}\big[\left(\tfrac{1}{2}\partial_w^2 f(W_j, t_j) R_j\right) \left(\tfrac{1}{2}\partial_w^2 f(W_k, t_j) R_k\right) \mid \mathcal{F}_j \big] \\
& = \mathrm{E}[\, R_j \mid \mathcal{F}_j] \; \frac{1}{4}\partial_w^2 f(W_j, t_j)\partial_w^2 f(W_k, t_j) R_k \; ,
\end{aligned}
$$

and the right side vanishes. The conditional expectation of a diagonal term is

$$
\begin{aligned}
\frac{1}{4}\mathrm{E}\Big[\left(\partial_w^2 f(W_j, t_j) R_j\right)^2 \mid \mathcal{F}_j \Big] &= \frac{1}{4}\left(\partial_w^2 f(W_j, t_j)\right)^2 \mathrm{E}\big[\, R_j^2 \mid \mathcal{F}_j \big] \\
&= \frac{1}{2}\left(\partial_w^2 f(W_j, t_j)\right)^2 \Delta t^2
\end{aligned}
$$

These calculations show that in $\mathrm{E}\big[\, S_{2,2}^2 \big]$, the diagonal terms, which are the only non-zero ones, sum to $\le C(b-a)\Delta t$.

The "almost surely" statement follows from the Borel Cantelli lemma, as last week. The abstract theorem is that if S_n is a family of random variables with

$$
\sum_{n=1}^{\infty} \mathrm{E}\big[\, S_n^2 \big] < \infty \; , \tag{9}
$$

eq:bc

then $S_n \to 0$ as $n \to \infty$ almost surely. This is because (9) implies that $S_n^2 \to 0$ as $n \to \infty$. If $S_n^2 \to 0$ then $S_n \to 0$ also. We know $S_n^2 \to 0$ almost surely because $S_n^2 \ge 0$ and if an infinite sum of positive numbers is convergent, then the terms go to zero. Our sum is convergent almost surely, so the sum is finite almost surely.

3 Backward equations

sec:be

Suppose $V(w)$ is a *running reward* function and consider

$$
f(w,t) = \mathrm{E}_{w,t}\left[\int_t^T V(W_s) ds \right] \; . \tag{10}
$$

eq:rr

As in the Introduction, this may be written in the equivalent form

$$f(W_t,t) = \mathrm{E}\left[\int_t^T V(W_s)ds \mid \mathcal{F}_t\right] \; . \tag{11}$$ $\boxed{\texttt{eq:rr2}}$

Ito's lemma gives

$$f(W_T,T) - f(W_t,t) = \int_t^T f_w(W_s,s)dW_s + \int_t^T \left(\frac{1}{2}f_{ww}(W_s,s) + f_t(W_s,s)\right)ds \; .$$

The definition (10) gives $f(W_T,T)=0$. Therefore, as in the Introduction,

$$f(W_t,t) = -\mathrm{E}\left[\int_t^T \left(\frac{1}{2}f_{ww}(W_s,s) + f_t(W_s,s)\right)ds \mid \mathcal{F}_t\right] \; .$$

We set the two expressions for f equal:

$$\mathrm{E}\left[\int_t^T V(W_s)ds \mid \mathcal{F}_t\right] = -\mathrm{E}\left[\int_t^T \left(\frac{1}{2}f_{ww}(W_s,s) + f_t(W_s,s)\right)ds \mid \mathcal{F}_t\right] \; .$$

The natural way to achieve this is to set the integrands equal to each other, which gives

$$\frac{1}{2}f_{ww}(w,s) + f_t(w,s) + V(w) = 0 \; . \tag{12}$$ $\boxed{\texttt{eq:berr}}$

The final condition for this PDE is $f(w,T)=0$. The PDE then determines the values $f(w,s)$ for $s<T$. Now that we have guessed the backward equation, we can show that it is right by Ito differentiation once more. If $f(w,s)$ satisfies the backward equation (12), then $f(W_t,t)$ satisfies (11).

Here is a slightly better way to say this. From ordinary calculus, we get

$$d\left(\int_t^T V(W_s)ds \mid \mathcal{F}_t\right) = -V(W_t)dt \; .$$

We pause to consider this. The stochastic process

$$X_t = \int_t^T V(W_s)ds$$

is a differentiable function of t. Its derivative with respect to t follows from the ordinary rules of calculus, the fundamental theorem in this case

$$\frac{dX}{dt}\int_t^T V(W_s)ds = -V(W_t) \; .$$

This is true for any continuous function W_t whether or not it is random. Conditioning on $\mathcal{F}_t$ just ties down the value of W_t. From Ito's lemma, any function $f(w,s)$ satisfies

$$\mathrm{E}[\,df(W_t,t) \mid \mathcal{F}_t] = \left(\frac{1}{2}f_{ww}(W_s,s) + f_t(W_s,s)\right)dt \; .$$

Taking expectations on both sides of (eq:rr2 11) gives

$$\left(\frac{1}{2}f_{ww}(W_s,s) + f_t(W_s,s)\right)dt = -V(W_t)dt \ ,$$

which is the backward equation (eq:berr 12).

Consider the specific example

$$f(w,t) = \mathrm{E}_{w,t}\left[\int_t^T W_s^2 dt\right] \ .$$

We could find the solution by direct calculations, since there is a simple formula $\mathrm{E}_{w,t}\left[W_s^2\right] = \mathrm{E}_{w,t}\left[W_t^2 + (W_s - W_t)^2\right] = w^2 + (s-t)$. Instead we use the *ansatz* method. Suppose the solution has the form $f(w,t) = A(t)w^2 + B(t)$. It is easy to plug into the backward equation

$$\frac{1}{2}f_{ww} + f_t + w^2 = 0$$

and get

$$2A + \dot{A}w^2 + \dot{B} + w^2 = 0 \ .$$

This gives $\dot{A} = -1$. Since $f(w,T) = 0$, we have $A(T) = 0$ and therefore $A(t) = T - t$. Next we have $\dot{B} = 2T - 2t$, so $B = 2Tt - t^2 + C$. The final condition $B(T) = 0$ gives $C = -T^2$. The simplified form is $B(t) = 2Tt - t^2 - T^2 = -(T-t)^2$. The solution is $f(w,t) = (T-t)w^2 - -(T-t)^2$.