Brownian motion and Ito calculus

Math 485

October 13, 2015

1 Brownian motion and Ito integral

1.1 Definition

A stochastic process W_t is a Brownian motion (abbreviated BM) if

   a. $W_0 = 0$

   b. $W_t - W_s$ is independent of $W_r, 0 \leq r \leq s$.

   c. $W_t - W_s$ has Normal(0, $t - s$) distribution.

   Remark:

   1. Property a can be changed to $W_0 = x$ which would be referred to as Brownian motion starting at x. Without specification, by default we refer to a Brownian motion starting at 0.

   2. Property b and c together are referred to as independent and stationary increment. Omiting the normal distribution, we can write it as: for any $s_1 < t1 \leq s_2 < t_2$ $W_{t_1} - W_{s_1}$ and $W_{t_2} - W_{s_2}$ are i.i.d. This can be generalized to n pair of points.

Example 1.1. *Compute, for $s < t$ $E(W_t)^2, E(W_s W_t), E(W_s W_t^2), E(W_t | W_s), E(W_t^2 | W_S)$.*

1.2 Some terminologies

1. We will say that an event E happens *almost surely* (abbreviated a.s.) (under some reference probability P) if it happens with probability 1 under P: $P(E) = 1$.

   2. A stochastic process can be viewed as a function of 2 variables: t and ω. Let X_t be a stochastic process (the ω variable is usually suppressed when it's implicitly understood that X_t also depends on ω). By *a path of X_t*, we understand it as fixing an ω and viewing $X(\cdot, \omega)$ as a function of t.

1.3   Some properties of Brownian motion

a. Almost surely, Brownian motion path is continous . That is

$$P(W_t \text{ is continuous in } t \text{ on } [0, T]) = 1.$$

b. Almost surely, Brownian motion has nowhere differentiable path. This property is one of the reasons we need to use Ito Calculus to study integration with respect to Brownian motion.

c. Let Δ be a finite partition of $t, T]$, that is a collection of $t = t_0 < t_1 < t_2 < \cdots < t_n = T$. Then for all partition Δ of $[t, T]$ we have

$$E \sum_i (W_{t_{i+1}} - W_{t_i})^2 = T - t.$$

We say the quadratic variation of Brownian motion on the interval $[t, T]$ is $T - t$. This is also one important feature of Brownian motion that leads to Ito Calculus.

Proof: See homework.

1.4   Integrating with respect to Brownian motion

As you have seen in the previous note, the underlying asset in continuous time has the structure

$$S_t = \int_0^t \alpha(u) d_u + \int_0^t \sigma(u) dW_u.$$

Thus we need to make sense out of the term $\int_0^t \sigma(u) dW_u$. Naively, one can define it like this: let Δ be a partition of $[0, t]$. By $\|\Delta\|$ we mean the mesh of this partition: $\max_i |t_{i+1} - t_i|$. Then following the Riemann integration technique, we can define

$$\int_0^t \sigma(u) dW_u := \lim_{\|\Delta\| \to 0} \sum_i \sigma(t_i)(W_{t_{i+1}} - W_{t_i}).$$

There are several issues that need to be considered.

1. In what sense do we understand the convergence in the RHS (if it's convergent at all)? The reason is both the LHS and the RHS are still functions of ω. So the most straightforward (!) way is to require the RHS to converge for all ω (or a.s.). This won't happen, because of the irregularity of BM. It turns out that we need to settle for some type of convergence on average (that is, when we average - take expectation

of the square - in ω, we see that the RHS converges). More specifically, we say X_n converges to Y in mean square if as $n \to \infty$

$$E[(X_n - Y)^2] \to 0.$$

In this way, there exists a RV, we call it $\int_0^T \sigma(u)dW_u$ so that $\sum_i \sigma(t_i)(W_{t_{i+1}} - W_{t_i})$ converges to $\int_0^T \sigma(u)dW_u$ as $\|\Delta\| \to 0$. This way of interpreting the convergence of the RHS is the first feature of Ito's integral.

2. Notice that the integrand σ is sampled at the left hand point of the partition: t_i. Recall that in Riemann integration, where you sample the integrand on the interval does not affect the value of the integral (you can use right hand, left hand or mid point). In integrating with respect to BM, it turns out that which sample point we use matters.

3. The choice of the left hand point is referred to as the Ito integral, which we use in math finance because of its connection to non-anticipating portfolio. The choice of the trapezoidal rule (using $\frac{1}{2}(\sigma(t_{i+1}) + \sigma(t_i))$) is referred to as the Stratonovich integral, and more popular among the physicists and engineers because of its closer connection to classical calculus and a "more natural" interpretation of convergence. ("More natural here means if we approximate the Brownian path with path of nicer poperty - say piecewise smooth - then the integrals against these paths converge to the Stratonovich integral.)

Example 1.2. *Let Δ be a partition of $[0, t]$. Then*

$$\sum_i 2W_{t_i}(W_{t_{i+1}} - W_{t_i})$$

converges to $W_t^2 - t$ in mean square as $\|\Delta\| \to 0$ while

$$2\sum_i \frac{1}{2}(W_{t_{i+1}} + W_{t_i})(W_{t_{i+1}} - W_{t_i})$$

converges to W_t^2 in mean square as $\|\Delta\| \to 0$.

Proof:
We show the first limit. The second is left as a homework exercise.
The key to proving the first limit is the identity

$$2W_{t_i}(W_{t_{i+1}} - W_{t_i}) = W_{t_{i+1}}^2 - W_{t_i}^2 - (W_{t_{i+1}} - W_{t_i})^2.$$

Thus

$$\sum_i 2W_{t_i}(W_{t_{i+1}} - W_{t_i}) = \sum_i W_{t_{i+1}}^2 - W_{t_i}^2 - \sum_i (W_{t_{i+1}} - W_{t_i})^2$$

$$= W_t^2 - \sum_i (W_{t_{i+1}} - W_{t_i})^2.$$

We will not show the details, but it can be showed without much difficulty that

$$\sum_i (W_{t_{i+1}} - W_{t_i})^2$$

converges to t in mean square (The proof is left as a homework exercise). Just keep in mind that it is not the same as showing

$$E(\sum_i (W_{t_{i+1}} - W_{t_i})^2) \to t,$$

which is a simpler calculation and indeed true for every partition Δ. What we need to show is

$$E\left\{ \left[\sum_i (W_{t_{i+1}} - W_{t_i})^2 - t \right]^2 \right\} \to 0,$$

as $\|\Delta\| \to 0$.

1.5 Some properties of the Ito integrals

1. With probability 1, $X_t = \int_0^t \alpha(u)dW_u$ is nowhere differentiable in t.
2.

$$E\left\{ \int_0^t \alpha(u)dW_u \right\} = 0.$$

$$E\left\{ \left[\int_0^t \alpha(u)dW_u \right]^2 \right\} = \int_0^t E(\alpha(u)^2)du.$$

3. If $\alpha(u)$ is **deterministic**, then $\int_0^t \alpha_u du$ has N(0, $\int_0^t \alpha(u)^2 du$) distribution.

Example 1.3. $\int_0^t udW_u$ has N(0, $\frac{t^3}{3}$) distribution.

Example 1.4. $W_t^2 - t = \int_0^t 2W_u dW_u$ *does NOT have a Normal distribution, because the integrand,* $2W_u$, *is NOT deterministic.*

We can still compute

$$E(W_t^2 - t) = t - t = 0,$$

which is consistent with

$$E(\int_0^t 2W_u dW_u) = 0.$$

Computing $E((W_t^2 - t)^2) = E(W_t^4 - 2tW_t^2 + t^2)$ *involves computing the fourth moment of a Normal distribution (we'll learn how to do that later with Ito's formula). There's an easier way to do it by computing the RHS:*

$$E\Big\{ \big[\int_0^t 2W_u dW_u \big]^2 \Big\} = \int_0^t 4E(W_u)^2 du = \int_0^t 4u du = 2t^2.$$

Thus we can deduce that

$$E(W_t^4) = 2t^2 - t^2 + 2tE(W_t^2) = 3t^2.$$

1.6   Ito integral versus Riemann integral

The convergence

$$\sum_i 2W_{t_i}(W_{t_{i+1}} - W_{t_i})$$

to $W_t^2 - t$ in mean square means that

$$\int_0^t 2W_u dW_u = W_t^2 - t.$$

Now suppose we have a process $x(t)$ that is differentiable in t with $x(0) = 0$, so that we can write

$$dx(t) = \dot{x}(t)dt,$$

where $\dot{x}(t)$ is the derivative of $x(t)$ wrt to t. (The above is just another way to write $\frac{dx(t)}{dt} = \dot{x}(t)$. Then we have, by the chain rule,

$$\int_0^t 2x(u)dx(u) = \int_0^t 2x(u)\dot{x}(u)du = x^2(t) - x^2(0) = x^2(t).$$

Thus Ito integral has a "correction" term versus its Riemannian counter part, in this case the $-t$ in the Ito integral. This can be thought of also as a consequence of the non-differentiability of the BM, so Ito integral cannot be done in the Riemannian way to begin with (the $\dot{W}(t) = \frac{dW(t)}{dt}$ does not exist).

1.7 Ito process and differential form

Let $\alpha(t, \omega)$ and $\sigma(t, \omega)$ be nice enough process. The discussion above have given the meaning to the process X_t defined as

$$X_t = x + \int_0^t \alpha(u)du + \int_0^t \sigma(u)dWu,$$

where the dt term is understood in the Riemannian sense and the dWt term is understood in the Ito sense. Any process X_t with the above structure is referred to as an Ito process. Note that we also have $X_0 = x$ in the above equation.

It is the convention in stochastic calculus that we write the above equation for X_t in differential form:

$$\begin{aligned} dX_t &= \alpha(t)dt + \sigma(t)dWt, \\ X_0 &= 0. \end{aligned}$$

Just keep in mind that the differential form has no more rigorous meaning than saying:

$$X_t = x + \int_0^t \alpha(u)du + \int_0^t \sigma(u)dWu.$$

So if this is your first time seeing the differential form, it is good to build the habit of automatically converting it to the integral form to give it meaning.

1.8 Explicit formula for Ito's integral

You may wonder if we have explicit formula for Ito integral calculation, like the example $\int_0^t W_u dW_u$. The answer is yes, for certain types of integrand, essentially like the classical calculus. Of course we do not want to go through the definition to learn what the answer is. Recall that we "compute" the integral in the classical case by guessing the anti-derivative, then prove that it is the true anti-derivative by differentiating this candidate.

The idea for the Ito integral is the same, except that we have the correction term, so it's not easy to guess precisely what the "anti-derivative" is - but we can get close. For example in $\int_0^t 2W_u dWu$ it is natural to guess it is W_t^2. The next step is to "differentiate" the candidate and find out what the correction term should be. However, we do not know what differentiation in the Ito's context means yet. It cannot be done like the classical derivative, because one can also show that the

process $Z_t = \int_0^t \sigma(u)dWu$ is nowhere differentiable in t. The Ito's formula gives the us the meaning in "differentiating" in this case, as well as rules to "differentiate" correctly with the right correction term.

2 Ito formula

2.1 Another perspective of differentiation

Let $f(t)$ be a deterministic function of t. What do we mean by the derivative of f wrt t: $f'(t)$? We can either use the classical definition, via the limit of the difference quotion, *or* we can define it as a function such that for all $s < t$

$$f(t) - f(s) = \int_s^t f'(u)du.$$

The only issue is such definition of $f'(t)$ needs to be unique (there might possibly be two different functions $f'(t)$ that would satisfy the above equation). As you may have expected from classical calculus, it turns out there is only one $f'(t)$ that satisfies the above equation.

In this way, one can develop classical calculus with the notion of the integral first, and define the derivative this way (which is the approach of the Lebesgue integral and the derivative associated with it).

As far as we're concerned, we've followed exactly this program with the Ito's integral. We developed what it means to have an Ito's integral, and now we ask what it means to differentiate in t. If you take the above definition as guidance, then we've already known the answer: the "derivative" of a stochastic process is dX_t. This statement contains nothing deep, it simply says:

$$X_t = \int_0^t dX_u,$$

which is true by partitioning the interval $[0, t]$ and check it the Riemannian way.

However, if X_t is an Ito process then we have something more to say: if

$$X_t = x + \int_0^t \alpha(u)du + \int_0^t \sigma(u)dWu,$$

then

$$dX_t \;=\; \alpha(t)dt + \sigma(t)dWt.$$

Note that if we divide both sides of the differential form with dt (a VERY formal operation, not to be taken in any rigorous sense) we have

$$\frac{dXt}{dt} = \alpha(t) + \sigma(t)\frac{dWt}{dt},$$

which is exactly the differentiation with respect to t that we have in mind. The problem, of course is the term $\frac{dWt}{dt}$ does not exist, so we have to use the differential form, and interpret the two types of integrals in two different sense.

So it turns out that differentiation in Ito sense is not that exciting (there is not much we can say). Yet there are still questions worth asking: if

$$dX_t = \alpha(t)dt + \sigma(t)dWt,$$

what can we say about dX_t^2, or better, $df(X_t)$ for a smooth function of $f(X_t)$? Asking this question is equivalent to formulating the chain rule in the classical case. Since if we have a differentiable function $x(t)$ in t, we have

$$df(x(t)) = f'(x(t))\dot{x}(t)dt.$$

In the case that $X(t)$ is an Ito process, $X(t)$ is *not* differentiable in t. Thus the above formula does not work. The chain rule for a function of an Ito process, i.e. to figure out $df(X_t)$ when X_t is an Ito proess, is referred to as the Ito's formula.

2.2 The Ito's formula

Theorem 2.1. *Let X_t be a Ito process:*

$$dX_t = \alpha(t)dt + \sigma(t)dWt.$$

Let $f(t, x)$ be once continously differentiable in t, twice continuously differentiable in x (abbreviated as $C^{1,2}$). Then

$$
\begin{aligned}
df(t, X_t) &= f_t(t, X_t)dt + f_x(t, X_t)dX_t + \frac{1}{2}f_{xx}(t, X_t)\sigma^2(t)dt \\
&= [f_t(t, X_t) + f_x(t, X_t)\alpha(t) + \frac{1}{2}\sigma^2(t)f_{xx}(t, X_t)]dt + f_x(t, X_t)\sigma(t)dWt.
\end{aligned}
$$

Remark: If we understood that the formula should be evaluated at (t, X_t) we can neglect the arguments and write the Ito's formula in a somewhat cleaner form:

$$
\begin{aligned}
df(t, X_t) &= f_t dt + f_x dX_t + \frac{1}{2}f_{xx}\sigma_t^2 dt \\
&= [f_t + f_x\alpha_t + \frac{1}{2}\sigma_t f_{xx}]dt + f_x\sigma_t dWt.
\end{aligned}
$$

Example 2.2. *Apply Ito's formula to W_t^2:*

$$dW_t^2 = 2W_t dW_t + dt$$

That is

$$W_t^2 - W_0^2 = \int_0^t 2W_u dW_u + t.$$

Or

$$\int_0^t 2W_u dW_u = W_t^2 - t,$$

recovering the formula we developed above.

Example 2.3. *Apply Ito's formula to tW_t^2:*

$$d(tW_t^2) = W_t^2 dt + 2tW_t dW_t + tdt.$$

That is

$$tW_t^2 = \int_0^t W_u^2 du + \int_0^t 2uW_u dW_u + \int_0^t udu.$$

Or

$$tW_t^2 - \frac{t^2}{2} = \int_0^t W_u^2 du + \int_0^t 2uW_u dW_u.$$

This is an example of the integration by parts formula in the Ito's context:

$$
\begin{aligned}
\int_0^t W_u^2 du &= tW_t^2 - \int_0^t ud(W_u^2) \\
&= tW_t^2 - \int_0^t 2udW_u - \int_0^t udu.
\end{aligned}
$$

2.3 Intuition of the Ito's formula

We won't give a rigorous proof of the Ito's formula. We'll only list here the steps to help you understand why such formula is intuitively correct. We'll do this by considering different forms of $f(t, X_t)$. Note that intuitively

$$df(t, X_t) \approx \Delta f(t, X_t) = f(t + h, X_{t+h}) - f(t, X_t),$$

for very small h. The key to Ito's formula is the second order Taylor expansion. Indeed Ito's formula can be looked at as Taylor's expansion with the dW_t term.

2.3.1 $f(t, X_t) = f(W_t)$

Apply Taylor's expansion

$$f(W_{t+h}) - f(W_t) = f'(W_t)(W_{t+h} - W_t) + \frac{1}{2}f''(W_{t+\varepsilon})(W_{t+h} - W_t)^2,$$

for some $0 \le \varepsilon \le h$.

Therefore, if Δ is a partition of $[0, t]$ such that $\|\Delta\| \le h$, we have

$$
\begin{aligned}
f(W_t) - f(W_0) &= \sum_i f(W_{t_{i+1}}) - f(W_{t_i}) \\
&= \sum_i f'(W_{t_i})(W_{t_i+h} - W_{t_i}) + \frac{1}{2}f''(W_{t_i+\varepsilon})(W_{t_i+h} - W_{t_i})^2
\end{aligned}
$$

Let $\|\Delta\| \to 0$, you'll see that

$$\sum_i f'(W_{t_i})(W_{t_i+h} - W_{t_i}) \to \int_0^t f(W_u)dW_u,$$

in mean square.

We have discussed how $\sum_i (W_{t_i+h} - W_{t_i})^2 \to t$ in mean square as $\|\Delta\| \to 0$. In a similar manner, you can believe that

$$\sum_i f''(W_{t_i+\varepsilon})(W_{t_i+h} - W_{t_i})^2 \to \int_0^t f''(W_t)dt,$$

in mean square. That is, we have showed

$$f(W_t) - f(W_0) = \int_0^t f'(W_u)dW_u + \frac{1}{2}\int_0^t f''(W_u)du.$$

2.3.2 $f(t, X_t) = f(t, W_t)$

This case is very similar to the above, except the Taylor's expansion becomes

$$
\begin{aligned}
f(t+h, W_{t+h}) - f(t, W_t) &= f_t(t, W_t)(t + h - h) + f_x(t, W_t)(W_{t+h} - W_t) \\
&+ \frac{1}{2}f_{xx}(t, W_t)(W_{t+h} - W_t)^2 + f_{tx}(t, W_t)h(W_{t+h} - W_t) \\
&+ \text{higher order term}
\end{aligned}
$$

where by higher order term we mean terms of order $h^k(W_{t+h} - W_t)^j$ such that $j + k \ge 3$.

Apply the partition of $[0, t]$ as before, and note that

$$\sum_i f_t(t_i, W_{t_i})h \quad \to \quad \int_0^t f(u, W_u)du,$$

$$\sum_i f_x(t_i, W_{t_i})(W_{t_i+h} - W_{t_i}) \quad \to \quad \int_0^t f_x(W_u)dW_u,$$

$$\sum_i f_{xx}(t_i, W_{t_i})(W_{t_i+h} - W_{t_i})^2 \quad \to \quad \int_0^t f_{xx}(W_t)dt,$$

$$\sum_i \text{higher order term} \quad \to \quad 0$$

in mean square as before. The only point to note is that the term

$$\sum_i f_{tx}(t_i, W_{t_i})(W_{t_i+h} - W_{t_i})h \quad \to \quad 0$$

as $h \to 0$.

The heuristic reason is this: since $\sum_i (W_{t_i+h} - W_{t_i})^2 \to t$ in mean square, you can believe that $W_{t_i+h} - W_{t_i}$ " $\to$ " $\sqrt{h}$ for small h. (This is not to be taken in any rigorous way). Thus, if f_{tx} is bounded by a constant C:

$$\left| \sum_i f_{tx}(t_i, W_{t_i})(W_{t_i+h} - W_{t_i})h \right| \leq C \sum_i h^{3/2} \leq C\frac{T}{h}h^{3/2} \to 0,$$

as $h \to 0$.

The above gives the following informal rule about product of differentials when we deal with Taylor expansion involving dWt.

2.3.3 Informal rule for product of differentials

We have

$$
\begin{aligned}
(dt)^2 &= 0; \\
dt dW_t &= 0; \\
(dW_t)^2 &= dt.
\end{aligned}
$$

More rigorously, the above means, for a partition Δ of $[0, T]$ and bounded f

$$\sum_i f(t_i)(t_{i+1} - t_i)^2 \quad \to \quad 0;$$

$$\sum_i f(t_i)(t_{i+1} - t_i)(W_{t_{i+1}} - W_{t_i}) \quad \to \quad 0;$$

$$\sum_i f(t_i)(W_{t_{i+1}} - W_{t_i})^2 \quad \to \quad \int_0^t f(s)ds$$

11

in mean square as $\|\Delta\| \to 0$.

With these informal rule, we can give an intuition about the Ito's formula for the general case that X_t is an Ito process.

2.3.4 $f(t, X_t), X_t$ an Ito process

By Taylor's expansion

$$
\begin{aligned}
f(t+h, X_{t+h}) - f(t, X_t) \;=\; & f_t(t, X_t)(t+h-t) + f_x(t, W_t)(X_{t+h} - X_t) \\
& + \frac{1}{2} f_{xx}(t, X_t)(X_{t+h} - X_t)^2 + f_{tx}(t, X_t)h(X_{t+h} - X_t) \\
& + \text{ higher order term}
\end{aligned}
$$

Note that

$$
X_{t+h} - X_t \approx \alpha(t)h + \sigma(t)(W_{t+h} - W_t).
$$

Rewriting the Taylor's expansion, replacing the difference term with differential term we have

$$
\begin{aligned}
df(t, X_t) \;=\; & f_t dt + f_x dX_t + \frac{1}{2} f_{xx}(\alpha_t^2 (dt)^2 \\
& + \sigma_t^2 (dWt)^2 + 2\alpha_t \sigma_t dt dW_t) + f_{tx} dt(\alpha_t dt + \sigma_t dW_t).
\end{aligned}
$$

Applying the informal rule of product of differentials, we get

$$
df(t, X_t) \;=\; f_t dt + f_x dX_t + \frac{1}{2} f_{xx} \sigma_t^2 dt,
$$

which is exactly the Ito's formula.

2.4 An application of Ito's formula

Ito's formula is fundamental in stochastic analysis as it provides the equivalence of the chain rule in this context. It also provides a nice tool for some computations, for example, the moments of a Normal RV as followed.

Example 2.4. *Let X have Normal$(0, \sigma^2)$ distribution. Compute $E(X^4)$.*

Ans: If we were to apply the usual formula

$$
E(X^4) = \int_{-\infty}^{\infty} x^4 f_X(x) dx,
$$

where $f_X(x)$ is the density of the Normal $(0, \sigma^2)$ it would be a tedious calculation.

Instead, let's compute $E(W_t^4)$ where W_t is a BM. If we can do this, then

$$E(X^4) = \sigma^4 E(Z^4) = \frac{\sigma^4}{t^2} E(W_t^4),$$

where Z is a standard Normal.

Now apply Ito's formula to W_t^4 we have

$$W_t^4 = W_0^4 + \int_0^t 4W_s^3 dW_s + \int_0^t 6W_s^2 ds.$$

The key is the expectation of $\int_0^t 4W_s^3 dW_s$ is 0. Thus taking expectation on both sides, we get

$$E(W_t^4) = 6 \int_0^t E(W_s)^2 ds = 6 \int_0^t s^2 ds = 3t^2.$$

Therefore $E(X^4) = 3\sigma^4$.

Remark: It is clear that $E(X^{2k+1})$, k an integer is always 0. So we're only interested in computing the even moments of X. And for that purpose, we can just focus on computing the even moment of W_t using the conversion trick described above.

Example 2.5. *Compute $E(W_t^{2k})$.*

Ans: Apply Ito's formula to W_t^{2k}:

$$W_t^{2k} = W_0^{2k} + \int_0^t 2kW_s^{2k-1} dW_s + \int_0^t k(2k-1)W_s^{2k-2} ds.$$

Taking expectations on both sides give

$$E(W_t^{2k}) = \int_0^t k(2k-1)E(W_s^{2k-2}) ds.$$

This provides a recursion formula to compute the even moments of W_t. For example, for $k = 3$

$$\begin{aligned} E(W_t^6) &= \int_0^t 3 \cdot 5 \cdot 3s^2 ds \\ &= 5 \cdot 3t^3. \end{aligned}$$

With this, we can compute for the case $k = 4$ for $E(W_t^8)$. In fact, it is not hard to guess that

$$E(W_t^8) = 7 \cdot 5 \cdot 3t^4,$$

and in general

$$E(W_t^{2k}) = (2k - 1)!!t^k,$$

where

$$(2k - 1)!! = (2k - 1)(2k - 3) \cdots 3 \cdot 1.$$