Chapter 18

Stochastic Integrals with the Wiener Process

Section 18.1 addresses an issue which came up in the last lecture, namely the martingale characterization of the Wiener process.

Section 18.2 gives a heuristic introduction to stochastic integrals, via Euler's method for approximating ordinary integrals.

Section 18.3 gives a rigorous construction for the integral of a function with respect to a Wiener process.

18.1 Martingale Characterization of the Wiener Process

Last time in lecture, I mentioned (without remembering much of the details) that there is a way of characterizing the Wiener process in terms of some martingale properties. Here it is.

Theorem 183 *If $M(t)$ is a continuous martingale, and $M^2(t) - t$ is also a martingale, then $M(t)$ is a Wiener process.*

There are some very clean proofs of this theorem[1] — but they require us to use stochastic calculus! Doob (1953, pp. 384ff) gives a proof which does not, however. The details of his proof are messy, but the basic idea is to get the central limit theorem to apply, using the martingale property of $M^2(t) - t$ to get the variance to grow linearly with time and to get independent increments, and then seeing that between any two times t_1 and t_2, we can fit arbitrarily many little increments so we can use the CLT.

We will return to this result as an illustration of the stochastic calculus.

[1] See especially Ethier and Kurtz (1986, Theorem 5.2.12, p. 290).

18.2 A Heuristic Introduction to Stochastic Integrals

Euler's method is perhaps the most basic method for numerically approximating integrals. If we want to evaluate $I(x) \equiv \int_a^b x(t)dt$, then we pick n intervals of time, with boundaries $a = t_0 < t_1 < \ldots t_n = b$, and set

$$I_n(x) = \sum_{i=1}^{n} x\left(t_{i-1}\right)\left(t_i - t_{i-1}\right)$$

Then $I_n(x) \rightarrow I(x)$, if x is well-behaved and the length of the largest interval $\rightarrow 0$. If we want to evaluate $\int_{t=a}^{t=b} x(t)dw$, where w is another function of t, the natural thing to do is to get the derivative of w, w', replace the integrand by $x(t)w'(t)$, and perform the integral with respect to t. The approximating sums are then

$$\sum_{i=1}^{n} x\left(t_{i-1}\right) w'\left(t_{i-1}\right)\left(t_i - t_{i-1}\right) \tag{18.1}$$

Alternately, we could, if $w(t)$ is nice enough, approximate the integral by

$$\sum_{i=1}^{n} x\left(t_{i-1}\right)\left(w\left(t_i\right) - w\left(t_{i-1}\right)\right) \tag{18.2}$$

(You may be more familiar with using Euler's method to solve ODEs, $dx/dt = f(x)$. Then one generally picks a Δt, and iterates

$$x(t + \Delta t) = x(t) + f(x)\Delta t \tag{18.3}$$

from the initial condition $x(t_0) = x_0$, and uses linear interpolation to get a continuous, almost-everywhere-differentiable curve. Remarkably enough, this converges on the actual solution as Δt shrinks (Arnol'd, 1973).)

Let's try to carry all this over to random functions of time $X(t)$ and $W(t)$. The integral $\int X(t)dt$ is generally not a problem — we just find a version of X with measurable sample paths (Section 8.2). $\int X(t)dW$ is also comprehensible if dW/dt exists (almost surely). Unfortunately, we've seen that this is not the case for the Wiener process, which (as you can tell from the W) is what we'd really like to use here. So we can't approximate the integral with a sum like Eq. 18.1. But there's nothing preventing us from using one like Eq. 18.2, since that only demands increments of W. So what we would like to say is that

$$\int_{t=a}^{t=b} X(t)dW \equiv \lim_{n\rightarrow\infty} \sum_{i=1}^{n} X\left(t_{i-1}\right)\left(W\left(t_i\right) - W\left(t_{i-1}\right)\right) \tag{18.4}$$

This is a crude-but-workable approach to numerically evaluating stochastic integrals, and apparently how the first stochastic integrals were defined, back in the 1920s. Notice that it is going to make the integral a *random variable*, i.e., a measurable function of ω. Notice also that I haven't said anything yet which should lead you to believe that the limit on the right-hand side exists, in any sense, or that it is independent of the choice of partitions $a = t_0 < t_1 < \ldots t_n$ b. The next section will attempt to rectify this.

(When it comes to the SDE $dX = f(X)dt + g(X)dW$, the counterpart of Eq. 18.3 is

$$X(t + \Delta t) = X(t) + f(X(t))\Delta t + g(X(t))\Delta W \tag{18.5}$$

where $\Delta W = W(t+\Delta t) - W(t)$, and again we use linear interpolation in between the points, starting from $X(t_0) = x_0$.)

18.3 Integrals with Respect to the Wiener Process

The drill by now should be familiar: first we define integrals of step functions, then we approximate more general classes of functions by these elementary functions. We need some preliminary technicalities.

Definition 184 (Progressive Process) *A continuous-parameter stochastic process X adapted to a filtration $\mathcal{G}$ is* progressively measurable *or* progressive *when $X(s, \omega)$, $0 \leq s \leq t$, is always measurable with respect to $\mathcal{B}_t \times \mathcal{G}_t$, where $\mathcal{B}_t$ is the Borel σ-field on $[0, t]$.*

If X has continuous sample paths, for instance, then it is progressive.

Definition 185 (Non-anticipating filtrations, processes) *Let W be a standard Wiener process, $\{\mathcal{F}_t\}$ the right-continuous completion of the natural filtration of W, and $\mathcal{G}$ any σ-field independent of $\{\mathcal{F}_t\}$. Then the* non-anticipating filtrations *are the ones of the form $\sigma(\mathcal{F}_t \cap \mathcal{G})$, $0 \leq t < \infty$. A stochastic process X is* non-anticipating *if it is adapted to some non-anticipating filtration.*

The idea of the definition is that if X is non-anticipating, we allow it to depend on the history of W, and possibly some extra, independent random stuff, but none of that extra information is of any use in predicting the future development of W, since it's independent.

Definition 186 (Elementary non-anticipating process) *A progressive, non-anticipating process X is* elementary *if there exist an increasing sequence of times t_i, starting at zero and tending to infinity, such that $X(t) = X(t_n)$ if $t \in [t_n, t_{n+1})$, i.e., if X is a step-function of time.*

Definition 187 (Square-integrable in the mean) *A random process X is* square-integrable from a to b *if $\mathbf{E}\left[\int_a^b X^2(t)dt\right]$ is finite.*

Notice that if X is bounded on $[a, b]$, in the sense that $|X(t)| \leq M$ with probability 1 for all $a \leq t \leq b$, then X is square-integrable from a to b.

Definition 188 (Itô integral of an elementary process) *If X is an elementary, progressive, non-anticipative process, square-integrable from a to b, then its* Itô integral from a to b *is*

$$\int_a^b X(t)dW \equiv \sum_{i\geq 0} X(t_i)\left(W\left(t_{i+1}\right) - W\left(t_i\right)\right) \tag{18.6}$$

where the t_i are as in Definition 186, truncated below by a and above by b.

Notice that this is basically a Riemann-Stieltjes integral. It's a random variable, but we don't have to worry about the existence of a limit. Now we set about approximating more general sorts of processes by elementary processes.

Lemma 189 *Suppose X is progressive, non-anticipative, bounded on $[a, b]$, and has continuous sample paths. Then there exist bounded elementary processes X_n, Itô-integrable on $[a, b]$, such that*

$$\lim_{n\to\infty} \mathbf{E}\left[\int_a^b (X - X_n)^2 dt\right] = 0 \tag{18.7}$$

Proof: Set

$$X_n(t) \equiv \sum_{i=0}^{\infty} X(t_i)\mathbf{1}_{[i/2^n,(i+1)/2^n)}(t) \tag{18.8}$$

This is clearly elementary, bounded and square-integrable on $[a, b]$. Moreover, for fixed ω, $\int_a^b (X(t,\omega) - X_n(t,\omega))^2 dt \to 0$, since $X(t, \omega)$ is continuous. So the expectation of the time-integral goes to zero by bounded convergence. □

Lemma 190 *Suppose X is progressive, non-anticipative, and bounded on $[a, b]$. Then there exist progressive, non-anticipative processes X_n which are bounded and continuous on $[a, b]$ such that*

$$\lim_{n\to\infty} \mathbf{E}\left[\int_a^b (X - X_n)^2 dt\right] = 0 \tag{18.9}$$

Proof: Let M be the bound on the absolute value of X. For each n, pick a probability density $f_n(t)$ on $\mathbb{R}$ whose support is confined to the interval $(-1/n, 0)$. Set

$$X_n(t) \equiv \int_0^t f_n(s-t)X(s)ds \tag{18.10}$$

$X_n(t)$ is then a sort of moving average of X, over the interval $(t-1/n, t)$. Clearly, it's continuous, bounded, progressively measurable, and non-anticipative. Moreover, for each ω,

$$\lim_{n\to\infty} \int_a^b (X_n(t,\omega) - X(t,\omega))^2 dt = 0 \tag{18.11}$$

because of the way we've set up f_n and X_n. By bounded convergence, Eq. 18.9 follows. □

Lemma 191 *Suppose X is progressive, non-anticipative, and square-integrable on $[a,b]$. Then there exist a sequence of random processes X_n which are progressive, non-anticipative and bounded on $[a,b]$, such that*

$$\lim_{n\to\infty} \mathbf{E}\left[\int_a^b (X - X_n)^2 dt\right] = 0 \tag{18.12}$$

PROOF: Set $X_n(t) = (-n \vee X(t)) \wedge n$. This has the desired properties, and the result follows from dominated (not bounded!) convergence. □

Lemma 192 *Suppose X is progressive, non-anticipative, and square-integrable on $[a,b]$. Then there exist a sequence of bounded elementary processes X_n such that*

$$\lim_{n\to\infty} \mathbf{E}\left[\int_a^b (X - X_n)^2 dt\right] = 0 \tag{18.13}$$

PROOF: Combine the preceding three lemmas. □

This lemma gets its force from the following result.

Lemma 193 *Suppose X is as in Definition 188, and in addition bounded on $[a,b]$. Then*

$$\mathbf{E}\left[\left(\int_a^b X(t)dW\right)^2\right] = \mathbf{E}\left[\int_a^b X^2(t)dt\right] \tag{18.14}$$

PROOF: Set $\Delta W_i = W(t_{i+1}) - W(t_i)$. Notice that ΔW_j is independent of $X(t_i)X(t_j)\Delta W_i$ if $i < j$, because of the non-anticipation properties of X. On the other hand, $\mathbf{E}\left[(\Delta W_i)^2\right] = t_{i+1} - t_i$, by the linear variance of the increments of W. So

$$\mathbf{E}\left[X(t_i)X(t_j)\Delta W_j \Delta W_i\right] = \mathbf{E}\left[X^2(t_i)\right](t_{i+1} - t_i)\delta_{ij} \tag{18.15}$$

Substituting Eq. 18.6 into the left-hand side of Eq. 18.14,

$$\mathbf{E}\left[\left(\int_a^b X(t)dW\right)^2\right] = \mathbf{E}\left[\sum_{i,j} X(t_i)X(t_j)\Delta W_j \Delta W_i\right] \tag{18.16}$$

$$= \sum_{i,j} \mathbf{E}\left[X(t_i)X(t_j)\Delta W_j \Delta W_i\right] \tag{18.17}$$

$$= \sum_i \mathbf{E}\left[X^2(t_i)\right](t_{i+1} - t_i) \tag{18.18}$$

$$= \mathbf{E}\left[\sum_i X^2(t_i)(t_{i+1} - t_i)\right] \tag{18.19}$$

$$= \mathbf{E}\left[\int_a^b X^2(t)dt\right] \tag{18.20}$$

where the last step uses the fact that X^2 must also be elementary. □

Theorem 194 *Let X and X_n be as in Lemma 192. Then the sequence $I_n(X) \equiv$*

$$\int_a^b X_n(t)dW \tag{18.21}$$

has a limit in L_2. Moreover, this limit is the same for any such approximating sequence X_n.

PROOF: For each X_n, $I_n(X(\omega))$ is an L_2 function of ω, by the fact that X_n is square-integrable and Lemma 193. Now, the X_n are converging on X, in the sense that

$$\mathbf{E}\left[\int_a^b (X(t) - X_n(t))^2 dt\right] \rightarrow 0$$

i.e., in an L_2 sense, but on the interval $[a, b]$ of the real line, and not on Ω. Nonetheless, because this is a convergent sequence, it must also be a Cauchy sequence, so, for every $\epsilon > 0$, there exists an n such that

$$\mathbf{E}\left[\int_a^b (X_{n+k}(t) - X_n(t))^2 dt\right] < \epsilon$$

for every positive k. Since X_n and X_{n+k} are both elementary processes, their difference is also elementary, and we can apply Lemma 193 to it. That is, for every $\epsilon > 0$, there is an n such that

$$\mathbf{E}\left[\left(\int_a^b (X_{n+k}(t) - X_n(t))dW\right)^2\right] < \epsilon$$

for all k. But this is to say that $I_n(X)$ is a Cauchy sequence in $L_2(\Omega)$, therefore it has a limit, which is also in $L_2(\Omega)$. If Y_n is another sequence of approximations of X by elementary processes, it is also a Cauchy sequence, and so must have the same limit. □

Definition 195 *Let X be progressive, non-anticipative and square-integrable on $[a, b]$. Then its* Itô integral *is*

$$\int_a^b X(t)dW \equiv \lim_n \int_a^b X_n(t)dW \tag{18.22}$$

taking the limit in L_2, with X_n as in Lemma 192. We will say that X is Itô-integrable on $[a, b]$.

Corollary 196 (The Itô isometry) *Eq. 18.14 holds for all Itô-integrable X.*

PROOF: Obvious from the approximation by elementary processes and Lemma 193.

18.4 Exercises

Exercise 18.1 (Basic Properties of the Itô Integral) *Prove the following, first for elementary Itô-integrable processes, and then in general.*

a

$$\int_a^c X(t)dW = \int_a^b X(t)dW + \int_b^c X(t)dW$$

almost surely.

b If c is any real constant, then, almost surely,

$$\int_a^b (cX(t) + Y(t))dW = c\int_a^b XdW + \int_a^b Y(t)dW$$

Exercise 18.2 (Martingale Properties of the Itô Integral) *Suppose X is Itô-integrable on $[a, b]$. Show that*

$$I_x(t) \equiv \int_a^t X(s)dW$$

$a \leq t \leq b$, is a martingale. What is $E[I_x(t)]$?

Exercise 18.3 (Continuity of the Itô Integral) *Show that $I_x(t)$ has continuous sample paths.*