Stochastic Calculus Notes, Lecture 5

Last modified October 24, 2002

1 Continuous time martingales

Continuous time martingales play a central role in calculations involving Brownian motion and diffusions. The Ito stochastic integral and the Ito solution of stochastic differential equations are formulated to insure that quantities without explicit "drift" terms are martingales. Martingales are easy to create and manipulate because of the continuous time version of Doob's stopping time theorem and because of martingale limit theorems that are beyond these lectures. Verifying that this or that stochastic process is a martingale often comes down to a Taylor series calculation, as we will see. Sometimes, this allows us to compute expectation values of interesting quantities.

1.1. Continuous time martingale: A continuous time stochastic process, Y_t is a martingale if $E[Y_{t'} \mid \mathcal{F}_t] = Y_t$ whenever $t' > t$. Most of what we said about discrete time martingales also holds in continuous time. The martingale property may be thought of as the statement that the increment, $\Delta Y = Y_{t'} - Y_t$, is uncorrelated with the present value, Y_t. This has the consequence that $\text{var}(Y_{t'}) = \text{var}(Y_t) + \text{var}(\Delta Y)$. This makes martingales the main example of random variables that are uncorrelated but not independent. Whether a process is a martingale depends on which filtration (expanding family of $\sigma-$algebras, $\mathcal{F}_t$) we use.

1.2. Example 1: This example is the first of three that illustrate Ito's lemma (see later). The process is $Y_t = X_t^2 - t$, where X_t is Brownian motion. We can verify the martingale property by checking that $E[X_{t'}^2 - X_t^2 \mid \mathcal{F}_t] = t' - t$. This is a simple calculation based on the increments property of Brownian motion: $X_{t'} = X_t + \Delta X$, where ΔX is a gaussian random variable independent of X_t with variance $t' - t$. Thus

$$E[X_{t'}^2 \mid \mathcal{F}_t] = E[\Delta X^2 \mid \mathcal{F}_t] + X_t^2 = t' - t + X_t^2 ,$$

which is the martingale property.

1.3. More on conditional expectation: This should go in an earlier lecture but is inserted here; better late than never. Suppose we have a family of probability measures, P_λ, indexed by a parameter, λ. Then we write $E_\lambda[F(\omega)]$ for the integral $\int_\Omega F(\omega) dP_\lambda(\omega)$. For example, if $\Omega = R$ and P_σ is Gaussian with mean zero and variance σ^2, then $E_\sigma[X^4] = 3\sigma^4$. (This is easy to check using integration.) Now suppose X_t is a Markov process, either a discrete process with some transition matrix, or a continuous time Brownian motion. The conditional expectations $f_t = E[F(X) \mid \mathcal{F}_t]$ are actually functions of X_t because of the Markov property. In particular, the conditional probability measures depend on a measurable way on $\mathcal{F}_t$ and therefore are functions of X_t. We denote these measures $P_{x,t}$.

1.4. $P_{x,t}$ for discrete Markov chains: In the discrete Markov chain case, the measure $P_{x,t}$ is over paths that start at time t with $X_t = x$ and then follow the usual transition probability rules given by the transition matrix, P. For example, if A is the event $X_t = x_1$, $X_{t+1} = x_2$, $X_{t+2} = x_3$, then $P_{x,t}(A) = 0$ if $x_1 \neq x$. But, if $x_1 = x$, then $P_{x,t}(A) = P_{x_1,x_2} \cdot P_{x_2,x_3}$.

1.5. Warning, $P_{x,t}$ is only for $\mathcal{H}_t$ events: Recall the notations: $\mathcal{F}_t$ generated by X_s for $0 \leq s \leq t$, $\mathcal{G}_t$ generated by X_t only, and $\mathcal{H}_t$ generated by X_s for $s \geq t$. The $P_{x,t}$ measures are really for $\mathcal{G}_t$ conditional expectation measures. The Markov property allows us to identify the conditional measures for conditional expectation with respect to $\mathcal{F}_t$ and $\mathcal{G}_t$ only for events in $\mathcal{H}_t$. For example, the definition of $P_{x,t}$ would not tell how to calculate $P_{x,t}(A)$ for the event $A = \{X_{t-1} \neq X_t\}$, and the conditional probabilities of this event in $\mathcal{F}_t$ and $\mathcal{G}_t$ are different. The definition of $P_{x,t}$ above only works for $\mathcal{H}_t$ measurable sets. In the general definition of measure has a, probability space Ω, a $\sigma-$algebra $\mathcal{F}$, and a probability function (measure) P. The $\sigma-$algebra for $P_{x,t}$ is $\mathcal{H}_t$.

1.6. $P_{x,t}$ for discrete Markov chains: The independent gaussian increments definition of Brownian motion works equally well for continuing paths with $X_t = x$ as for paths with $X_0 = 0$. Because $P_{x,t}$ only gives probabilities about X_s for $s \geq t$, $P_{x,t}(A)$ is only defined for $A \in \mathcal{H}_t$.

1.7. Examples through PDE: Any solution of the backwards equation provides an example of a martingale. If $f(x,t)$ satisfies the backward equation with "final data" $f(x,T) = V(x)$, then $f(X_t,t)$ is a martingale. This is a consequence of the representation $f(x,t) = E_{x,t}[V(X_T)]$. It is an expression of the "tower property" of conditional expectation: losing information in stages is the same as losing the same information all at once: if $\mathcal{F}_1 \subset \mathcal{F}_2$ then $E[E[F \mid \mathcal{F}_2] \mid \mathcal{F}_1] = E[F \mid \mathcal{F}_1]$. In particular, if $t' > t$, then $\mathcal{F}_t \subset \mathcal{F}_{t'}$, so $Y_{t'} = E[V(X_T) \mid \mathcal{F}_{t'}$ has $E[Y_{t'} \mid \mathcal{F}_t] = Y_t$.

1.8. Doob's stopping time theorem. Let τ be a bounded stopping time. That means that there is some T so that $\tau \leq t$. We might add (but need not) the phrase "almost surely". Doob's stopping time theorem states that if Y_t is a martingale, then $\tilde{Y}_t = Y_{\min(t,\tau)}$ is also a martingale. This $\tilde{Y}_t$ is the original process Y_t stopped at time τ. That is, $\tilde{Y}_t = Y_t$ for $t \leq \tau$ and $\tilde{Y}_t = Y_\tau$ for $t > \tau$. This is really just a restatement of the other statement that $E[Y_\tau] = Y_0$.

Here is a trick based on these ideas. Suppose $f(x,t) = E_{x,t}[V(X_T)]$, $f(x_0,t) = 0$, and $tau = \min(t \mid f(X_t,t)0 =)$. Then

2 Integrals involving Brownian motion

The par of stochastic calculus often called the Ito calculus is a differential and integral calculus appropriate for Brownian motion and related stochastic processes. In ordinary calculus is made powerful by the relation between differentiation and integration that we call the "fundamental theorem of calculus": we can calculate the value of an integral if we know enough about differentiation. The corresponding ideas in stochastic calculus are called Ito's lemma and Dynkin's formula.

Before discussing the fundamental theorem, many calculus classes spend time calculating a few integrals directly from the definition as a limit of Riemann sums, $x^2/2 = \int_0^x tdt$ being a common example. After this, these cumbersome calculations help the student understand the definition of integration but are happily forgotten after the fundamental theorem makes them unnecessary. This section of notes only does things the hard way. The easy way is coming.

Integrals involving Brownian motion arise in many applications. One of the most famous is the "Feynman-Kac" formula. In financial applications, we get ingegrals of Brownian motion or related stochastic processes when we value Asian style options (options on the average price of an asset over a given time period) or compute the present value of a future payment with uncertain interest rates.

2.1. The integral of Brownian motion: Consider the random variable, where X_t continues to be standard Brownian motion,

$$Y = \int_0^T X_t dt \; . \tag{1}$$

Because X_t is a continuous function of t, there is no technical trouble defining the integral. Riemann sums converge. For this, choose n and define $\Delta t = T/n$ and $t_k = k\Delta t$ as usual. The Riemann sum approximation is

$$Y^{(n)} = \Delta t \sum_{k=0}^{n-1} X_{t_k} \; , \tag{2}$$

and $Y^{(n)} \to Y$ as $n \to \infty$. Each of the $Y^{(n)}$ is Gaussian because the right side of (2) is a sum of components of a multivariate normal. The mean is zero, as each term on the right has mean zero. The variance is computed below. The integral (1) is not an Ito integral. The Ito integral, dX_t in place of dt, which makes the definition more subtle. Here we have the integral *of* X_t, not with respect to X_t.

2.2. The variance of Y_T: We will start the hard way, computing the variance from (2) and letting $\Delta t \to 0$. The trick is to use two summation variables $Y^{(n)} = \Delta t \sum_{k=0}^{n-1} X_{t_k}$ and $Y^{(n)} = \Delta t \sum_{j=0}^{n-1} X_{t_j}$. Then we can write

$$E[Y^{(n)2}] \;\; = \;\; E[Y^{(n)} \cdot Y^{(n)}]$$

$$
\begin{aligned}
&= E\left[\left(\Delta t \sum_{k=0}^{n-1} X_{t_k}\right) \cdot \left(\Delta t \sum_{j=0}^{n-1} X_{t_j}\right)\right] \\
&= \Delta t^2 \sum_{jk} E[X_{t_k} X_{t_j}] \; .
\end{aligned}
$$

If we now let $\Delta t \to 0$, we get

$$
E[Y_T^2] = \int_{s=0}^{T} \int_{t=0}^{T} E[X_t X_s] ds dt \; . \qquad (3)
$$

We can find the needed $E[X_t X_s]$ if $s > t$ by writing $X_s \; X_t + \Delta X$ with ΔX independent of X_t, which gives $E[X_t(X_t + \Delta X)] = E[X_t X_t] = t$. Clearly $E[X_t X_s] = s$ if $s < t$. Altogether, $E[X_t X_s] = \min(t,s)$, which is a famous formula. This now gives

$$
E[Y_T^2] = \int_{s=0}^{T} \int_{t=0}^{T} E[X_t X_s] ds dt = \frac{1}{3} T^3 \; .
$$

There is a simpler and equally rigorous way to get this. Write $Y = \int_{s=0}^{T} X_s ds$ and $\int_{t=0}^{T} X_t dt$ so that again

$$
\begin{aligned}
E[Y_T^2] &= E\left[\int_{s=0}^{T} X_s ds \cdot \int_{t=0}^{T} X_t dt\right] \\
&= \int_{s=0}^{T} \int_{t=0}^{T} E[X_s X_t] dt ds ; .
\end{aligned}
$$

From the first to second lines I just interchanged the order of integration. After all, $E[\cdot]$ just represents integration over a probability space. The change in integration order could be written abstractly as

$$
\int_{\omega \in \Omega} \int_{s \in [0,t]} \int_{t \in [0,t]} F(\omega, s, t) dt ds dP(\omega) = \int_{s \in [0,t]} \int_{t \in [0,t]} \int_{\omega \in \Omega} F(\omega, s, t) dP(\omega) ds dt \; .
$$

In our situation, P is Brownian motion measure (Weiner measure) and $F = X_s X_t$.

2.3. Example 2, the X_t^3 martingale: As we will see later, many martingales are constructed from integrals involving Brownian motion. A simple one is

$$
Y_t = X_t^3 - 3 \int_0^t X_s ds \; .
$$

To check the martingale property, choose $t' > t$ and, for $s > t$, write $X_s = X_t + \Delta X_s$. Then

$$
E\left[\int_0^{t'} X_s ds \mid \mathcal{F}_t\right] = E\left[\int_0^{t} X_s ds + \int_t^{t'} X_s ds \mid \mathcal{F}_t\right]
$$

$$
\begin{aligned}
&= E\left[\int_0^t X_s ds \mid \mathcal{F}_t\right] + E\left[\int_t^{t'} (X_t + \Delta X_s)\, ds \mid \mathcal{F}_t\right] \\
&= \int_0^t X_s ds + (t'-t)X_t \ .
\end{aligned}
$$

In the last line we use the fact that $X_s \in \mathcal{F}_t$ when $s < t$, and, of course $X_t \in \mathcal{F}_t$, and that $E[\Delta X_s \mid \mathcal{F}_t] = 0$, which is part of the independent increments property. For the X_t^3 part, we have

$$
\begin{aligned}
E\left[(X_t + \Delta X_{t'})^3 \mid \mathcal{F}_t\right] &= E\left[X_t^3 + 3X_t^2\Delta X_{t'} + 3X_t\Delta X_{t'}^2 + \Delta X_{t'}^3 \mid \mathcal{F}_t\right] \\
&= X_t^3 + 3X_t^2 \cdot 0 + 3X_t E[\Delta X_{t'}^2 \mid \mathcal{F}_t] + 0 \\
&= X_t^3 + 3(t'-t)X_t \ .
\end{aligned}
$$

In the last line we used the independent increments property to get $E[\Delta X_{t'} \mid \mathcal{F}_t] = 0$, and the formula for the variance of the increment to get $E[\Delta X_{t'}^2 \mid \mathcal{F}_t] = t' - t$. This verifies that $E[Y_{t'} \mid \mathcal{F}_t] = Y_t$, which is the martingale property.

2.4. Backward equations for expected values of integrals: Many other integrals involving Brownian motion arise in applications and may be "solved" using backward equations. One example is $Y = \int_0^T V(X_s)ds$, which represents the total accumulated $V(X)$ over a Brownian motion path. Also, Y/T is the average of Y over the path. As before, this integral is a standard Riemann integral because the integrand, $V(X_s)$, is a continuous function of s. We can calculate $E[Y]$, using the more general function

$$
f(x,t) = E_{x,t}\left[\int_t^T V(X_s)ds\right] \ . \tag{4}
$$

Referring to paragraphs 1.3 through 1.6, we see that this can also be written

$$
f_t = E\left[\int_t^T V(X_s) \mid \mathcal{F}_t\right] \ ,
$$

with f_t being a function of X_t and $f_t(X_t) = f(X_t, t)$. The backward equation is

$$
\partial_t f + \frac{1}{2}\partial_x^2 f + V(x,t) = 0 \ , \tag{5}
$$

with final conditions $f(x,T) = 0$. Even if we are interested in the single number $E[Y] = f(0,0)$, the backward equation approach asks us to compute the entire function $f(x,t)$ by solving (5).

The derivation of the backward equation (5) is similar to the one we used before for the backward equation for $E_{x,t}[V(X_T)]$ We use Taylor series and the "tower property" to calculate how f changes over a small time increment, Δt. First, we have

$$
\int_t^T V(X_s)ds = \int_t^{t+\Delta t} V(X_s)ds + \int_{t+\Delta t}^T V(X_s)ds \ .
$$

The first integral on the right has the value $V(x)\Delta t + o(\Delta t)$. We write $o(\Delta t)$ for a quantity that is smaller than Δt in the sense that $o(\Delta t)/\Delta t \to 0$ as $\Delta t \to 0$ (we will shortly divide by Δt, take the limit $\Delta t \to 0$, and neglect all $o(\Delta t)$ terms.). The second term is measurable in $\mathcal{H}_{t+\Delta t}$, and

$$E\left[\int_{t+\Delta t}^{T} V(X_s)ds \mid \mathcal{F}_{t+\Delta t}\right] = f_{t+\Delta t}(X_{t+\Delta t}) \; .$$

Writing $X_{t+\Delta t} = X_t + \Delta X$, we use the tower property with $\mathcal{F}_t \subset \mathcal{F}_{t+\Delta t}$ to get

$$E\left[\int_{t+\Delta t}^{T} V(X_s)ds \mid \mathcal{F}_t\right] = E\left[f_{t+\Delta t}(X_t + \Delta X) \mid \mathcal{F}_t\right] \; .$$

Now we use Taylor expansions as before for the right side:

$$f(x+\Delta X, t+\Delta t) = f(x,t) + \Delta t \partial_t f(x,t) + \Delta X \partial_x f(x,t) + \frac{1}{2}\Delta X^2 \partial_x^2 f(x,t) + o(\Delta t) \; .$$

Therefore, changing notation slightly and taking the expected values,

$$E_{x,t}\left[f(x+\Delta X, t+\Delta t\right] = f(x,t) + \Delta t \partial_t f(x,t) + \frac{1}{2}\Delta t \partial_x^2 f(x,t) + o(\Delta t) \; .$$

The tower property gives, since $f_{t+\Delta t}$ is an $\mathcal{F}_{t+\Delta t}$ expectation,

$$E\left[\int_{t+\Delta t}^{T} V(X_s)ds \mid \mathcal{F}_t\right] = E\left[f_{t+\Delta t}(X_t + \Delta X) \mid \mathcal{F}_t\right] \; .$$

Combing these gives

$$f(x,t) = \Delta t V(x) + f(x,t) + \Delta t \partial_t f(x,t) + \frac{1}{2}\Delta t \partial_x^2 f(x,t) + o(\Delta t) \; .$$

Now just cancel $f(x,t)$ from both sides and let $\Delta t \to 0$ to get the promised equation (5).

2.5. Application of PDE: Most commonly, we cannot evaluate either the expected value (4) or the solution of the partial differential equation (PDE) (5). How does the PDE represent progress toward evaluating f? One way is by suggesting a completely different computational procedure. If we work only from the definition (4), we would use Monte Carlo for numerical evaluation. Monte Carlo is notoriously slow and inaccurate. There are several techniques for finding the solution of a PDE that avoid Monte Carlo, including finite difference methods, finite element methods, spectral methods, and trees. When such deterministic methods are practical, they generally are more reliable, more accurate, and faster. In financial applications, we are often able to find PDEs for quantities that have no simple Monte Carlo probabilistic definition. Many such examples are related to optimization problems: maximizing an expected return or minimizing uncertainty with dynamic trading strategies in a randomly evolving market. The Black Scholes evaluation of the value of an American style option is a well known example.

2.6. The Feynman Kac formula: The function

$$f(x,t) = E_{x,t}\left[e^{\int_t^T V(X_s)ds}\right] \tag{6}$$

satisfies the backward equation

$$\partial_t f + \frac{1}{2}\partial_x^2 f + V(x)f = 0 \ . \tag{7}$$

When someone refers to the "Feynman Kac formula", they usually are referring to the fact that (6) is a formula for the solution of the PDE (7). In our work, the situation mostly will be reversed. We use the PDE (7) to get information about the quantity defined by (6) or even just about the process X_t. A precurser to the formula, the "Feynman integral" solution to the Schrödinger equation, was given by physicist Richard Feynman (*Surely You're Joking, Mr. Feynman*). The Feynman integral, while well defined, is not an integral in the sense of measure theory. The colorful probabilist Marc Kac (pronounced "Katz") discovered that an actual integral over Wiener measure (6) gives the solution of (7).

We can verify that (6) satisfies (7) more or less as in the preceding paragraph. We note that

$$\begin{aligned}
&\exp\left\{\int_t^{t+\Delta t} V(X_s)ds + \int_{t+\Delta t}^T V(X_s)ds\right\} \\
&\quad = \exp\left\{\int_t^{t+\Delta t} V(X_s)ds\right\} \cdot \exp\left\{\int_{t+\Delta t}^T V(X_s)ds\right\} \\
&\quad = (1 + \Delta t V(X_t) + o(\Delta t)) \cdot \exp\left\{\int_{t+\Delta t}^T V(X_s)ds\right\}
\end{aligned}$$

The expectation of the rigth side with respect to $\mathcal{F}_{t+\Delta t}$ is

$$(1 + \Delta t V(X_t) + o(\Delta t))\, f_t(X_t + \Delta X) \ .$$

When we now take expectation with respect to $\mathcal{F}_t$, which amounts to averaging over ΔX, using Taylor expansion of f about $f(x,t)$ as before, we get (7).

2.7. Application of Feynman Kac: The problem of evaluating

$$f = E\left[\exp\left(\int_0^T V(X_t)dt\right)\right]$$

arises in many situations. In finance, f could represent the present value of a payment in the future subject to unknown fluctuating interest rates. The PDE (7) provides a possible way to evaluate $f = f(0,0)$, either analytically or numerically.