Module 12
Stochastic Differential Equations

1 Brownian Motion

Before we can really get into the details of stochastic differential equations, we need to discuss Brownian motion. Brownian motion is motion that is, intuitively, entirely random. It is the culmination (or the addition) of several random variables that are normally distributed, independent, and have 0 mean. Brownian motion is often referred to as a Wiener process. The paper I have asked you to read denotes the value of Brownian motion (or the Wiener process) at time t as $W(t)$. Here are some properties of the Wiener process that are listed in the paper:

1. $W(0) = 0$. The Brownian motion, at time 0, is 0. This makes some computations later in the module a little bit easier.

2. For $0 \leq s < t \leq T$, $W(t) - W(s) \sim N(0, \text{Var} = t - s)$. This simply states that the difference between two values of the Wiener process at different times is 0 mean (on average, the process isn't going to go anywhere), and the variance of the difference grows with the time difference This makes intuitive sense since it follows that the longer the time difference, the more time the process has to change, and the more variable that change becomes.

3. For $0 \leq s < t < u < v \leq T$, the measurents $W(t) - W(s)$ and $W(v) - W(u)$ are independent. In other words, the changes in the process at separate times are independent of one another.

The article considers a Wiener process from time 0 to time T, and it considers $N+1$ values of the Wiener process within this time inteveral, $[0, T]$. $W_0 = W(0)$ is the first value in time and $W_N = W(N)$ is the last value in time, and these values progress as

$$W_{j+1} = W_j + dW_j \quad j = 0, 1, 2, \ldots, N-1 \tag{1}$$

where dW_j is the jump (or change) in the process from time j to time $j+1$. From property (3) listed above, we know that

$$dW_0, dW_1, dW_2, \ldots, dW_N$$

are all independent, and from property (2) listed above, we know that if there are $N+1$ evenly spaced values of W, then the spacing between the time values is

$$\delta t = T/N \quad \& \quad dW_j \sim N(0, \text{Var} = \delta t).$$

All of these properties will become quite useful later in the module.

2 Stochastic Differential Equations and Solving Them

2.1 A Linear Stochastic Differential Equation

With what we just learned about Brownian Motion, we can now get into stochastic differential equations. In this module, we want to solve (or approximately solve) a stochastic differential equation (SDE). The article I asked you to read is good, but in my opinion, the ideas are all over the place and not very well organized. This is my attempt to consolidate and possibly better explain what is going on.

Consider the linear stochastic differential equation

$$dX(t) = f(X(t))dt + g(X(t))dW(t). \tag{2}$$

The intuition behind Equation (2) is relatively intuitive. Think about (2) as identifying/characterizing the change in a stock price. This (instantaneous) change is $dX(t)$. Equation (2) states that this change is related to its current price

$(f(X(t))dt)$ plus some random (white noise) factor $(g(X(t))dW(t))$ that might represent something like a weather event or some other unforseen event that may affect the stock-market. So the way to think of Equation (2) is

$$\underbrace{dX(t)}_{\text{Change in price}} = \underbrace{f(X(t))dt}_{\text{Some function of current price}} + \underbrace{g(X(t))dW(t)}_{\text{Random event times function of current price}}. \tag{3}$$

Solving the stochastic differential equation in (2) would give us what we want, which is $X(t)$, or the value of the stock as a function of time. Analytically, this is solved as

$$X(t) = X(0) + \int_0^t f(X(s))ds + \int_0^t g(X(s))dW(s) \tag{4}$$

and this module focuses on how to calculate/approximate the solution given in Equation (4). The following subsection in this section focuses on two different ways to calculate integrals such as the ones given above.

2.2 Calculating a Stochastic Integral

This section focuses on calculating values similar to the last term in Equation (4).

We all know how to solve an equation such as

$$\int_0^t f(s)ds \tag{5}$$

where $f(s)$ is deterministic. But what does an integral mean when it is summing over a stochastic (i.e., non-deterministic) function like in Equation (2)? The article I asked you to read motivates this quite well and here is my attempt to summarize/explain what they cover. The integral in (5) can be approximated with

$$\int_0^t f(s)ds \approx \sum_{j=0}^{N-1} f(s_j)(s_{j+1} - s_j) \quad \text{where} \quad s_N = t \;\&\; s_0 = 0. \tag{6}$$

Note that the approximation in Equation (6) evaluates f at the endpoint of the intervals $(s_0, s_1), (s_1, s_2), \ldots, (s_{N-1}, s_N)$. An alternative (and perhaps superior) way to approximate the integral in (5) is by evaluating f at the midpoint of the intervals $(s_0, s_1), (s_1, s_2), \ldots, (s_{N-1}, s_N)$, which would give

$$\int_0^t f(s)ds \approx \sum_{j=1}^{N-1} f\left(\frac{s_{j+1} + s_j}{2}\right)(s_{j+1} - s_j). \tag{7}$$

With these approximations for a deterministic integral in mind, now consider the stochastic integral in Equation (2). The article I asked you to read specifically focuses on a stochastic integral of the form

$$\int_0^t W(s)dW(s), \tag{8}$$

where $W(s)$ and $dW(s)$ are the values of a Wiener process (as discussed in the first section). There are two ways to approximate an integral like the one in (8). One way resembles the deterministic approach in (6), and the other resembles the deterministic approach in (7). The first approach is the Ito integral and the second is the Stratonovic integral.

2.2.1 The Ito Integral

The Ito integral approximates the integral in Equation (8) with

$$\int_0^t W(s)dW(s) \approx \sum_{j=0}^{N-1} W(s_j)\left(W(s_{j+1}) - W(s_j)\right).$$

Expanding the right side of the above equation out, we get

$$\begin{aligned}
\int_0^t W(s)dW(s) &\approx \sum_{j=0}^{N-1} W(s_j)\left(W(s_{j+1})-W(s_j)\right) \\
&= \frac{1}{2}\left(2W(s_j)W(s_{j+1})-2W(s_j)^2\right) \\
&= \frac{1}{2}\sum_{j=0}^{N-1}\left[W(s_{j+1})^2-W(s_j)^2-W(s_{j+1})^2+2W(s_j)W(s_{j+1})-W(s_j)^2\right] \\
&= \frac{1}{2}\sum_{j=0}^{N-1}\left[W(s_{j+1})^2-W(s_j)^2-\left(W(s_{j+1})-W(s_j)\right)^2\right]. \qquad (9)
\end{aligned}$$

Consider the term $\left(W(s_{j+1})-W(s_j)\right)^2$ above. Observe that

$$E\left[W(s_{j+1})-W(s_j)\right]^2 = \underbrace{\mathrm{Var}\left[W(s_{j+1})-W(s_j)\right]}_{=\delta t}+\underbrace{\left[E\left(W(s_{j+1})-W(s_j)\right)\right]^2}_{=0},$$

making

$$\sum_{j=0}^{N-1}\left(W(s_{j+1})-W(s_j)\right)^2 \approx E\left[W(s_{j+1})-W(s_j)\right]^2 = N\times\delta t = T.$$

With this in mind, the value in (9) goes to

$$\begin{aligned}
&\frac{1}{2}\sum_{j=0}^{N-1}\left[W(s_{j+1})^2-W(s_j)^2-\left(W(s_{j+1})-W(s_j)\right)^2\right] \\
\approx\ & \frac{1}{2}\left[W(s_1)^2-W(s_0)^2+W(s_2)^2-W(s_1)^2+\cdots+W(s_n)^2-W(s_{n-1})^2\right] \\
& -\frac{1}{2}\sum_{j=0}^{N-1}E\left[W(s_{j+1})-W(s_j)\right]^2 \\
=\ & \frac{1}{2}W(T)^2-\frac{1}{2}T.
\end{aligned}$$

2.2.2 The Stratonovich Integral

The Stratonovich integral approximates the integral in Equation (8) with

$$\int_0^t W(s)dW(s) \approx \sum_{j=0}^{N-1}\left(\frac{W(s_{j+1})+W(s_j)}{2}+\Delta Z_j\right)\left(W(s_{j+1})-W(s_j)\right), \qquad (10)$$

where $\Delta Z_j \sim N\left(0, \mathrm{Var}=\delta t/4\right)$. This is a slight adjustment to how we would expect the Stratonovich integral to approximate the integral in Equation (8). Given the discrete ingtegral approximations discussed at the beginning of Section 2.2, one would thaink that the appropriate approximation would be

$$\int_0^t W(s)dW(s) \approx \sum_{j=0}^{N-1}\left(\frac{W(s_{j+1})+W(s_j)}{2}\right)\left(W(s_{j+1})-W(s_j)\right). \qquad (11)$$

The value of ΔZ is added because the term $\left(W(s_{j+1})+W(s_j)\right)/2$ approximates a value of W at the midpoint of the interval. But at the midpoint, we would expect additional error, and this error has mean 0, and standard devaition of the length of the interval divided by 2 (std dev $=\sqrt{\delta t}/2 \Longrightarrow \mathrm{Var}=\delta t/4$). The Stratonovich integral thus has the addition of ΔZ_j.

Using math similar to that illustrated for the Ito integral, the Stratonovich integral in Equation (10) collapses to

$$\frac{1}{2}\left(W(T)^2 - W(0)^2\right) + \sum_{j=0}^{N-1} \Delta Z_j \left(W(s_{j+1}) - W(s_j)\right). \tag{12}$$

The right side of the sum in Equation (12) has an expected value of 0, and with that in mind, the Stratonovich integral can be approximated with

$$\int_0^t W(s)dW(s) \approx \frac{1}{2}W(T)^2.$$

2.3 The Euler-Maruyama Method

Now I'll talk about a numerical method used to solve the differential equation in (2). It is simple and called the Euler-Maruyama method. It's essientially a discretized version of the stochastic differential equation. Let $W_0 = 0$, $\delta t = T/N$, $W_{k+1} = W_k + dW_k$ where $dW_k \sim N(0, \text{Var} = \delta t)$, and then the values of X_j can be generated as

$$\begin{aligned} \underbrace{X_j - X_{j-1}}_{\approx dX(t)} &= \underbrace{f(X_{j-1})\delta t}_{\approx f(X(t))dt} + \underbrace{g(X_{j-1})dW_k}_{\approx g(X(t))dW(t)} \\ \Longrightarrow X_j &= X_{j-1} + f(X_{j-1})\delta t + g(X_{j-1})dW_k. \end{aligned}$$

3 The Convergence of Solutions for Stochastic Differential Equations

In Section 2.3 of the notes, we discussed the Euler-Maruyama (from now on, it will be called the "EM") method - a numerical method/way of solving a stochastic differential equation. An obvious question related to this solution is "How well does it work?" The article discusses two possible ways to evaluate the quality of a solution: strong and weak convergence.

A solution to a stochastic differential equation converges "strongly" (and has order γ) if there exists a constant C such that

$$\mathbb{E}\left|X_j - X_j^{\text{soln}}\right| \leq C\left(\Delta t\right)^\gamma.$$

A solution to a stochastic differential equation converges "weakly" (and has order γ) if there exists a constant C such that

$$\left|\mathbb{E}\left(X_j\right) - \mathbb{E}\left(X_j^{\text{soln}}\right)\right| \leq C\left(\Delta t\right)^\gamma.$$

In my opinion, the easiest way to understand these two types of convergence is to compare them. Strong convergence examines the expected value of the differences, while weak convergence examines the difference in expected values. To understand the distinction between these two, consider the following two situations:

Situation 1 Let $X_1, X_2, \ldots, X_n$ be independent random variables such that $X_j \sim N\left(0, \text{Var} = \frac{1}{j}\right)$ and let $Y_1, Y_2, \ldots, Y_n$ be independent random variables such that $Y_j \sim N\left(\frac{1}{j}, \text{Var} = \frac{1}{j}\right)$. If these are independent sequences, then

$$X_j - Y_j \sim N\left(-\frac{1}{j}, \text{Var} = \frac{2}{j}\right),$$

and

$$\mathbb{E}\left|X_j - Y_j\right| = \frac{1}{j} \longrightarrow 0.$$

In this case the expected difference of X_j and Y_j converges to 0 (similar to strong convergence) because the values of X_j and Y_j become more and more similar.

Situation 2 Now consider the sequence of random variables $W_1, W_2, \ldots, W_n$. Assume they are independent and identically distributed such that

$$W_j = \begin{cases} 1 & \text{with probability } 1/2 \\ -1 & \text{with probability } 1/2 \end{cases}.$$

And consider the variables $Z_1, Z_2, \ldots, Z_n$ such that

$$Z_j \sim N\left(0, \mathrm{Var} = \frac{1}{n}\right).$$

In this case, it is easy to see that $|W_j - Z_j|$ gets stuck at 1. The expected values of W_j and Z_j are identical, however, making the difference in the expected values 0, i.e.,

$$|\mathbb{E}(W_j) - \mathbb{E}(Z_j)| = 0.$$

This is similar to weak convergence.