Stochastic Calculus and Black-Scholes pricing

Ameya Vikrama Singh

July 2023

Introduction

Here, we define the common terminology:

- $(\Omega, \mathcal{B}, \mathbb{P})$ is our probability space, where Ω is a set called the sample space.
- $\mathcal{B}$ is a set of subsets of Ω, which is closed under **countable union** and **complementation**, called a Borel σ-algebra. A subalgebra is a subset of $\mathcal{B}$ which is a σ-algebra. The set of subalgebras is $\mathcal{B}(\Omega)$
- $\mathbb{P} : B \to \mathbb{R}$ is a probability measure, with $P(\Omega) = 1$ and satisfying the rule of disjoint addition
- A **filtration** $\mathcal{F}$ is defined as a sequence of σ-algebras such that each member of the sequence is a subset of it's successor.
- A continuous time filtration $\mathcal{F}(t) : \mathbb{R} \to \mathcal{B}(\Omega)$ satisfies $\mathcal{F}(t_1) \subseteq \mathcal{F}(t_2)$ if $t_1 \leq t_2$.
- The σ-algebra associated with a random variable X, $\sigma(X)$, is the set $\{\{A : X(A) = \alpha\} : \alpha \in \mathbb{R}\}$
- A random variable is said to be measurable by σ-algebra B iff $\sigma(X) \subseteq B$
- The usual notions of independence and conditioning follow from probability theory.

Adapted Stochastic Processes

A stochastic process $X(t)$ is said to be adapted to a filtration $\mathcal{F}(t)$, if the random variable $X(t)$ is measurable by $\mathcal{F}(t)$ for all t. The idea of an adapted process is extremely important for **Ito Calculus**, where the Ito integral is only defined for adapted processes.

An adpated stochastic process is said to be a **martingale**, if $\mathbb{E}[X(t)|X(t_1)] = X(t_1)$ for all $t \geq t_1$, i.e the martingale has no tendency to rise or fall at any point of time.

An adapted stochastic process is said to satisfy the **Markov property**, if for all functions f, there exists function g such that $\mathbb{E}[f(X(t))|X(t_1)] = g(X(t_1))$ for all $t \geq t_1$. Note that it is not necessary for every martingale to be a Markov process, since the Markov property needs to be true for all functions f. Further, every Markov process need not be a martingale, since the function g need not be the identity function. The name *martingale* comes from a betting strategy, for an interesting reference see this

Brownian Motion - The Wiener Process

Consider the stochastic process satisfying the following: For all $0 = t_0 < t_1 < t_2 < \cdots t_m$, the increments $W(t_{i+1}) - W(t_i)$ are independent and normally distributed, with mean 0 and variance $t_{i+1} - t_i$

$$W(t_{i+1}) - W(t_i) = \mathcal{N}(0, t_{i+1} - t_i)$$

This Brownian motion extends the idea of the discrete time coin-toss process, and perfectly attains the normal limit of the binomial process.

The Covariance

Since the individual increments are all normally distributed, the sequence of random variables $W(t_1), W(t_2) \cdots W(t_n)$ for $t_1 < t_2 < \cdots t_n$ are distributed normally in n dimensions. The convariance is given by

$$Cov(W(t_1), W(t_2)) = \mathbb{E}[W(t_1)W(t_2)] = t_1$$

Thus, we have the covariance matrix:

$$C = \begin{bmatrix} t_1 & t_1 & t_1 & \cdots & t_1 \\ t_1 & t_2 & t_2 & \cdots & t_2 \\ t_1 & t_2 & t_3 & \cdots & t_3 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ t_1 & t_2 & t_3 & \cdots & t_n \end{bmatrix}$$

This covariance matrix is true for all sequences $t_1 < t_2 < \cdots t_n$ and serves as an equivalent characterization of Brownian Motion

Martingale Property

The Brownian motion satisfies the Martingale property, since the increments have expected value 0, and are independent of conditioning.

Quadratic Variation

The Brownian motion process is highly discontinuous in increments. This is because it represents a large scale view of the coin toss process, hence it involves jumps at even the smallest time scale. To quantify this discontinuous derivative, stochastic calculus uses the notion of quadratic variation. This non-zero quadratic variation, i.e the fact that $dWdW$ **accumulates a finite integral over time**, represents a volatility term in those stochastic processes that depend on an underlying Brownian motion.

For discrete time stochastic processes, quadratic variation can be defined straightforward with the step size:

$$Q = \sum_{i=0}^{n} [W(t_{i+1}) - W(t_i)]^2$$

In the case of continuous time processes like the Wiener process, we must take a limit as the width of partition tends to 0, similar to the case of the Riemann Integral. It is clear that for functions with a continuous derivative will accumulate 0 quadratic variation. For stochastic processes, the quadratic variation is a random variable, and we consider the limit of this random variable's distribution.

Here, we provide a simple proof of the following theorem. For more details, see Stochastic Calculus for Finance II by Steven Shreve.

Theorem (Almost Sure Quadratic Variation)**.** *The quadratic variation of the Wiener process, denoted [W, W](T) converges almost surely to T, for all T.*

Proof. First, note that $E[[W,W](T)] = \sum_{|\Pi| \to 0} E[[W(t_{i+1} - W(t_i)]^2] = \sum_{|\Pi \to 0|} Var(W(t_{i+1} - W(t_i)) = T$
Next, we observe that the variance of this random variable is bounded linearly by the partition width, and hence tends to 0 as the width tends to 0. $\square$

An alternative and important way to represent the above fact can be seen in the following way:

Consider $Y_{i+1} = \dfrac{W(t_{i+1}) - W(t_i)}{\sqrt{t_{i+1} - t_i}}$. This is a standard normal random variable. The quadratic variation for a uniform partition of the time interval can be represented in terms of Y_is as $\sum \dfrac{Y_i^2 T}{n}$, which converges to T, as per the Law of Large Numbers, since $E[Y^2] = 1$.

An informal way to write it is the following equation, which defines the differential dW:

$$dWdW = dt$$

The Ito Integral

For an interesting exposition of the Ito integral, see this.

The Ito integral of a stochastic process $f(t)$ with respect to Brownian motion is defined similar to a Riemann sum, except since the Wiener process is volatile and can change even during a partition, we cannot really "approximate" the area with rectangles exactly. In fact, the volatile motion with quadratic variation leads to an additional term occuring in the Ito integral.

$$I(t) = \lim_{|\Pi| \to 0} \sum f(t_i)[W(t_{i+1}) - W(t_i)]$$

Martingale Property

It can be demonstrated as follows that the Ito integral is a martingale, since every partition satsifies the martingale property:

Let s and t be such that $s < t$ and s, t are in different partitions, let $s \in [t_l, t_{l+1})$ and $t \in [t_k, t_{k+1})$

We have

$$I(t) = \sum_{0}^{l-1} f(t_i)[W(t_{i+1}) - W(t_i)] + f(t_l)[W(t_{l+1}) - W(t_l)] + \sum_{l+1}^{k-1} + f(t_k)[W(t) - W(t_k)]$$

Since all times before s are known, the first term is constant in expectation at s. The second term has expectation:

$$E[f(t_l)[W(t_{l+1}) - W(t_l)]] = f(t_l)(W(s) - W(t_l))$$

The sum of these expectations is $I(s)$. It can be easily shown that the remaining sum has expectation 0 at s, since the increments of Brownian motion are normal random variables. Thus, the martingale property is satisfied.

Quadratic Variation

The quadratic variation of the Ito integral can be shown to be

$$[I, I](t) = \int_0^t f^2(u)du$$

since, each partition subinterval gains $f^2(t)dt$ quadratic variation, the Riemann integral can be used in RHS. Note that the QV is a random variable. In simpler notation we write

$$dIdI = f^2 dW dW = f^2 dt$$

Ito-Doeblin Formula

The Ito Doeblin formula for Brownian motion generalizes the chain rule of differentation to stochastic calculus, as follows:

Consider a real function f, and the random variable $f(W)$, then

$$\begin{aligned} df(W) &= f(W + dW) - f(W) \\ &= f'(W)dW + \frac{1}{2} f"(W) dW dW \\ &= f'(W)dW + \frac{1}{2} f"(W) dt \end{aligned}$$

This represents the essence of the Ito-Doeblin formula for Brownian motion

Black-Scholes Pricing

Now we can go ahead and derive the Black Scholes equation which models the price of options based on certain underlying assumptions:

The Black-Scholes equation is derived under the following assumptions:

1. The stock price follows a geometric Brownian motion:

$$dS = \mu S \, dt + \sigma S \, dW$$

   where:

   - S is the stock price,
   - μ is the drift rate (expected return),
   - σ is the volatility of the stock,
   - dW is a Wiener process increment (normal distribution with mean 0 and variance dt).

2. No dividends are paid during the option's life.
3. No transaction costs or taxes.
4. The risk-free rate r is constant and known.
5. Markets are frictionless (no arbitrage opportunities).
6. The option is European, meaning it can only be exercised at expiration.

Portfolio Construction

Consider a European call option $C(S, t)$ on a stock S, where t is the time. Construct a portfolio Π by holding:

- Δ units of stock S,
- and one short position in the option C (i.e., we sell the option).

The portfolio value is:

$$\Pi = \Delta S - C(S, t)$$

Change in Portfolio Value

The change in the portfolio value over a small time dt is:

$$d\Pi = \Delta \, dS - \frac{\partial C}{\partial t} \, dt - \frac{\partial C}{\partial S} \, dS - \frac{1}{2} \frac{\partial^2 C}{\partial S^2} \, (dS)^2$$

Using Itô's Lemma to express dC:

$$dC = \frac{\partial C}{\partial t} \, dt + \frac{\partial C}{\partial S} \, dS + \frac{1}{2} \frac{\partial^2 C}{\partial S^2} \, \sigma^2 S^2 \, dt$$

Thus, the change in the portfolio value $d\Pi$ becomes:

$$d\Pi = \Delta dS - \left(\frac{\partial C}{\partial t} \, dt + \frac{\partial C}{\partial S} \, dS + \frac{1}{2} \frac{\partial^2 C}{\partial S^2} \sigma^2 S^2 \, dt \right)$$

Simplifying, we get:

$$d\Pi = \left(\Delta - \frac{\partial C}{\partial S} \right) dS - \frac{\partial C}{\partial t} dt - \frac{1}{2} \frac{\partial^2 C}{\partial S^2} \sigma^2 S^2 \, dt$$

Eliminate Risk (Delta-Hedging)

To eliminate risk, we set the coefficient of dS to zero:

$$\Delta = \frac{\partial C}{\partial S}$$

This makes the portfolio riskless, implying that its return should equal the risk-free rate r. Therefore:

$$d\Pi = r\Pi\, dt = r\left(\frac{\partial C}{\partial S} S - C\right) dt$$

Substituting into the equation for $d\Pi$:

$$r\left(\frac{\partial C}{\partial S} S - C\right) dt = \frac{\partial C}{\partial t}\, dt + \frac{1}{2}\frac{\partial^2 C}{\partial S^2}\sigma^2 S^2\, dt$$

Dividing by dt and rearranging terms, we get the Black-Scholes PDE:

$$\frac{\partial C}{\partial t} + rS\frac{\partial C}{\partial S} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 C}{\partial S^2} - rC = 0$$