STOCHASTIC CALCULUS FOR ARBITRAGE FREE PRICING WITH STOCHASTIC VOLATILITY

DANIEL LI

Abstract. We provide an introduction to Ito calculus and stochastic processes, specifically Brownian motion, in order to rigorously define integration with respect to Brownian motion. We then introduce the Feynman-Kac formula to provide intuition for basic pricing and arbitrage. We use this formula to derive the Black-Scholes model for pricing options using arbitrage free assumptions. Modifying the initial assumptions of constant volatility, we then derive the Heston model using stochastic volatility.

Contents

1. Probability Spaces and Random Variables

We will begin with an introduction to certain probabilistic concepts in order to better understand stochastic processes in the following sections.

Definition 1.1. (Sample Space). A sample space Ω is the set of all outcomes of some procedure. After a procedure has happened, exactly one outcome must have occurred. Additionally, each outcome has an assigned probability to occur.

Definition 1.2. (Distribution and Density Functions). The distribution function $F : \mathbb{R} \to [0, 1]$ of a random variable on its sample space Ω is defined as

$$F(x) = \mathcal{P}(X \le x).$$

Additionally, if X is a continuous random variable and if there exists a function f such that

$$F(x) = \int_{-\infty}^{x} f(y)\, dy,$$

then f is the density function of X.

Definition 1.3. (Expectation). The expectation of a continuous random variable X is defined as

$$\mathbb{E}[X] = \int_{\Omega} X \, d\mathbb{P}$$

and if X has density f, then

$$\mathbb{E}[X] = \int_{-\infty}^{\infty} x f(x) \, dx.$$

If a discrete random variable has no density function with values $x_1, x_2, ..., x_n$ then the expectation is given by

$$\mathbb{E}[X] = \sum_{i=1}^{n} x_i \mathbb{P}(X = x_i).$$

Definition 1.4. (Variance). The variance of a random variable X is

$$\mathrm{Var}[X] = \mathbb{E}[(X - \mathbb{E}[X])^2] = \mathbb{E}[X^2] - \mathbb{E}[X]^2$$

Definition 1.5. (Central Limit Theorem). If $X_1, X_2, ..$ are independent, identically distributed random variables with generic $\mathbb{E}[X_i] = \mu$ and $\mathrm{Var}[X_i] = \sigma^2 < \infty$ and letting

$$Z_n = \frac{(X_1 + ... + X_n) - n\mu}{\sigma\sqrt{n}},$$

then as $n \to \infty$, the distribution of Z_n approaches the standard normal distribution.

2. Stochastic Processes

We now introduce some concepts that are required to understand stochastic calculus.

Definition 2.1. (Stochastic Process). A collection of random variables $\{X_t | t \in \mathbb{R}^+\}$ indexed by time t is called a stochastic process. We can view this object in two ways. First, we can consider it as a random variable whose value is a function from T to $\mathbb{R}$. We can also consider for each t, a random variable with correlations between the values at different times.

Definition 2.2. (Filtration). Letting $X_1, X_2, ..$ be a sequence of random variables, the associated filtration is the collection $\{\mathcal{F}_n\}$ with $\mathcal{F}_n$ denoting the information in $X_1, ..., X_n$. One assumption in the definition is that no information is lost.

Before we introduce what a martingale is, it is useful to know some properties of a sequence of random variables.

Proposition 2.3. *Let $X_1, X_2, ...$ be a sequence of random variables with filtration $\mathcal{F}_n$. The conditional expectation $\mathbb{E}[Y|\mathcal{F}_n]$ satisfies the following:*

- *If Y is $\mathcal{F}_n$ measurable, then $\mathbb{E}[Y|\mathcal{F}_n] = Y$.*
- *If E is an $\mathcal{F}_n$ measurable event, then*
$$\mathbb{E}[E[Y|\mathcal{F}_n]] = \mathbb{E}[Y]$$
- ***Linearity.*** *If X, Y are random variables with constants a,b, then*
$$\mathbb{E}[aY + bZ|\mathcal{F}_n] = a\mathbb{E}[X|\mathcal{F}_n] + b\mathbb{E}[Y|\mathcal{F}_n]$$
- ***Tower Property.*** *If $n < m$, then*
$$\mathbb{E}[E[Y|\mathcal{F}_m]|\mathcal{F}_n] = \mathbb{E}[Y|\mathcal{F}_n]$$

- *If X is an $\mathcal{F}_n$ measurable random variable, then with respect to $\mathcal{F}_n$, Z acts like a constant,*
$$\mathbb{E}[YZ|\mathcal{F}_n] = Z\mathbb{E}[Y|\mathcal{F}_n]$$

Definition 2.4. (Martingale). A collection of random variables $M_0, M_1, ...$ is a martingale with respect to filtration $\{\mathcal{F}_n\}$ if the following holds:

- For each t, M_t is a $\{\mathcal{F}_n\}$ measurable random variable with $\mathbb{E}[|M_n|] < \infty$
- if $n < m$, then
$$\mathbb{E}[M_m|\mathcal{F}_n] = M_n \text{ or restated } \mathbb{E}[M_m - M_n|\mathcal{F}_m] = 0.$$

Intuitively, a martingale is a model of a fair game. Viewing it in this way implies that regardless of what has happened so far, the expected future winnings is 0. In particular, it is a fair game.

3. Brownian Motion

In this section, we will provide an intuitive description of Brownian Motion and explore its formal definition. Recall that stochastic processes are collections of random variables with defined time intervals. We can similarly view Brownian Motion as a stochastic process where the defined time intervals become very small or equivalently as the limit of a random walk. We must be careful when taking this limit to preserve the magnitude of steps relative to time.

Suppose $X_1, X_2, ...$ are independent random variables with $P\{X_j = 1\} = P\{X_j = -1\} = 1/2$, and let
$$S_n = X_1 + ... + X_n$$
be the corresponding random walk. To simulate a discrete random walk, we can choose the time increment to be $\Delta t = 1$ and the space increment to be $\Delta x = 1$. Instead, let us set $\Delta t = 1/N$ with N large. Then, at every time step $\Delta t, 2\Delta t, ...,$, the value of S_n either jumps up or down Δx. If we view the process at time $t = 1 = N\Delta t$, the value of the process is
$$W_1^{(N)} = \Delta x(X_1 + ... + X_N).$$
We then find Δx such that $\text{Var}[W_1^{(N)}] = 1$. Since we know that
$$\text{Var}[\Delta x(X_1 + ... + X_N)] = (\Delta x)^2[\text{Var}(X_1 + ... + X_N)] = (\Delta x)^2 N,$$
then
$$\Delta x = \sqrt{1/N} = \sqrt{\Delta t}.$$
If N becomes large, then the distribution of
$$\frac{X_1 + ... + X_N}{\sqrt{N}}$$
is roughly normal with mean 0 and variance 1.

This exploration of the limit of a random walk gives us basic intuition for what standard Brownian motion looks like. Below, we write the formal definition of Brownian motion, which can also be thought of as continuous random motion.

Definition 3.1. (Brownian Motion). A stochastic process B_t is called a Brownian motion with drift m and variance σ^2 starting at the origin if it satisfies the following conditions:

- $B_0 = 0$

- If $s < t$, then the variable $B_t - B_s$ is normally distributed with mean $m(t-s)$ and variance $\sigma^2(t-s)$.
- If $s < t$, then the variable $B_t - B_s$ is independent of the value of B_r for $r \leq s$.
- The function $t \to B_t$ is a continuous function of t with probability one.

B_t is called standard Brownian motion if $m = 0$ and $\sigma^2 = 1$.

The proof that such a process actually exists is quite long and not necessary for the purposes of understanding the uses of Brownian Motion.

4. Ito Calculus

Before exploring stochastic calculus, it is useful to review ordinary differential equations. Consider the equation $df(t) = C(t, f(t))dt$, or equivalently, $\frac{df}{dt} = f'(t) = C(t, f(t))$. If we are given $f(0) = x_0$, then a solution can be given by

$$f(t) = x_0 + \int C(s, f(s))\, ds.$$

In, stochastic calculus, we attempt to make sense of equations of the form

$$dX_t = m(t, X_t)dt + \sigma(t, X_t)dB_t,$$

where B_t is standard Brownian motion. A solution to this equation is

$$X_t = X_0 + \int_0^t m(s, X_s)\, ds + \int_0^t \sigma(s, X_s)dB_s.$$

The *Ito integral* will give a precise meaning to the last term of this equation.

The analog of a step function for the stochastic integral is a simple process.

Definition 4.1. (Simple Process). A process A_t is a simple process if there exists times

$$0 = t_0 < t_1 < ... < t_n < t_{n+1} < \infty$$

and random variables $Y_j, j = 0, 1, ..., n$ that are $\mathcal{F}_{t_j}$ measurable such that

$$A_t = Y_j, t_j \leq t < t_{j+1}.$$

Then since Y_j is $\mathcal{F}_{t_j}$ measurable, A_t is $\mathcal{F}_{t_j}$ measurable.

We also assume that $\mathbb{E}[Y_j^2] < \infty$ for each j. If A_t is a simple process, then we define $Z_t = \int_0^t A_s\, dB_s$ by

$$Z_{t_j} = \sum_{i=0}^{j-1} Y_i[B_{t_{i+1}} - B_{t_i}]$$

and more generally,

$$Z_t = Z_{t_j} + Y_j[B_t - B_{t_j}] if t_j \leq t \leq t_{j+1},$$

$$\int_r^t A_s\, dB_s = Z_t - Z_r.$$

Based on these definitions, we note four important properties of the stochastic integral of simple processes.

Proposition 4.2. *Suppose A_t, C_t are simple processes with $Z_t = \int_0^t A_s\, dB_s$, and let B_t be a standard Brownian motion.*

- ***Linearity.*** *Let a, b be constants. Then $aA_t + bC_t$ is also a simple process with*
$$\int_0^t (aA_s + bC_s)\, dB_s = a\int_0^t A_s\, dB_s + b\int_0^t C_s\, dB_s$$
- ***Martingale property.*** *The process Z_t is a martingale with respect to $\{\mathcal{F}_t\}$*
- ***Variance rule.***
$$\mathrm{Var}[Z_t] = \mathbb{E}[Z_t^2] = \int_0^t \mathbb{E}[A_s^2]\, ds$$
- ***Continuity.*** *The function $t \to Z_t$ is a continuous function with probability one.*

We seek to generalize the integral of A_t first to bounded, continuous paths, then account for unbounded paths, and lastly piece-wise continuous paths. In order to do so, we first state the following approximation result.

Lemma 4.3. *Let A_t be a process with continuous paths, adapted to a filtration $\{\mathcal{F}_t\}$. Then suppose there exists $C < \infty$ such that $|A_t| \le C$ for all t with probability one. Then there exists a sequence of simple processes $A_t^{(n)}$ such that for all t,*
$$\lim_{n\to\infty} \int_0^t \mathbb{E}[|A_s - A_s^{(n)}|^2] ds = 0.$$
and also for all $n, t, |A_t^{(2)}| \le C$.

Definition 4.4. (Ito Integral for Bounded Process with Continuous Paths). Let A_t be a bounded process with continuous paths adapted to filtration $\{\mathcal{F}_t\}$. Then there exists a sequence of simple processes $A_t^{(n)}$ such that
$$\lim_{n\to\infty} \int_0^t \mathbb{E}\left[\left|A_s - A_s^{(n)}\right|^2\right] ds = 0$$
Then we can define
$$Z_t = \lim_{n\to\infty} \int_0^t A_s^{(n)}\, dB_s$$

Definition 4.5. (Ito Integral for Unbounded Processes with Continuous Paths). Let A_t be an unbounded process with continuous paths adapted to filtration $\{\mathcal{F}_t\}$. Let $T_n = \inf\{t : |A_t| = n\}$ for all $n = 0, 1, ... < \infty$. Then $A_t^{(n)} = A_{\min(t,T_n)}$ is a sequence of bounded and continuous processes with corresponding Ito integrals $Z_t^{(n)} = \int_0^t A_s^{(n)}\, dB_s$. Then let
$$Z_t = \lim_{n\to\infty} Z_t^{(n)}.$$

For unbounded processes, continuity, linearity, and the variance rule still hold, but it is possible that $Var[Z_t] = \infty$. The Martingale, however, may not hold as A_s may grow to infinity. That being said, the integral is still a *local martingale*, which is defined below. Intuitively, a local martingale is a martingale that is stopped before the values get too large.

Definition 4.6. (Local Martingale). A continuous process M_t adapted to the filtration $\{\mathcal{F}_t\}$ is a *local martingale* on $[0, T]$ if there exists a sequence of stopping times

$$\tau_1 \leq \tau_2 \leq ...$$

such that

$$\lim_{j \to \infty} \tau_j = T$$

with probability one, and for each j,

$$M_t^{(j)} = M_{\min(t, \tau_j)}$$

is a martingale.

Using the above generalizations, we can now understand stochastic differential equations and their integral forms like

$$dX_t = m(t, X_t)dt + \sigma(t, X_t)dB_t$$

$$X_t = X_0 + \int_0^t m(s, X_s)\, ds + \int_0^t \sigma(s, X_s) dB_s.$$

in a well-defined manner.

We will now use what we have described so far to understand Ito's Lemma. Ito's Lemma is the stochastic analog of the chain rule. In using Taylor expansions, stochastic equations require more terms due to their non-differentiability. The proof is not necessary for understanding its applications in the further sections. We call a function C^k if it has k continuous derivatives.

Theorem 4.7. (Ito's Formula) *Let f(t,x) be C^1 in t and C^2 in x. Then, letting B_t be a standard Brownian motion,*

$$f(t, B_t) = f(0, B_0) + \int_0^t \partial_x f(s, B_s)\, dB_s + \int_0^t \left[\partial_s f(s, B_s) + \frac{1}{2}\partial_{xx} f(s, B_s)\right] ds,$$

or in differential form,

$$df(t, B_t) = \partial_x f(t, B_t) dB_t + \left[\partial_t f(t, B_t) + \frac{1}{2}\partial_{xx} f(t, B_t)\right] dt.$$

Before exploring the applications of stochastic calculus in finance, it is important to understand geometric Brownian motion.

Definition 4.8. (Geometric Brownian Motion). Letting B_t be a standard Brownian motion, a process X_t is a geometric Brownian motion with volatility σ and drift m if it satisfies the SDE

$$dX_t = mX_t dt + \sigma X_t dB_t = X_t[mdt + \sigma dB_t]$$

The exact expression for geometric Brownian motion is

$$X_t = X_0 \exp\left\{\left(m - \frac{\sigma^2}{2}\right)t + \sigma B_t\right\},$$

although it is much more useful in terms of its SDE.

5. Feynman-Kac

One of the basic uses of Ito's Lemma is deriving the Black-Scholes formula for options pricing. However, we first seek to derive the Feynman-Kac equation, which can be used as another pricing model. First, to establish some terminology, a European call option is the right to buy an underlying asset at a strike price K, at a specified time T. A European put option is the right to sell an underlying asset at a strike price K, at a specified time T. In contrast to European options, American options may be exercised at any time up till the specified time of expiry T.

Note that an option gives the right (but not the obligation) to buy or sell the stock at the strike price, thus capping an option owner's loss to the cost of the option. Options may thus be worth more than the strike minus the current price of the underlying asset. Now suppose that an underlying stock's price follows the geometric Brownian motion

$$dS_t = mS_t dt + \sigma S_t dB_t \tag{5.1}$$

We will begin our exploration of options pricing using the most naive approach—by pricing it according to its inflation adjusted expected profit at that exact moment. Given the definition of an option, we will only exercise it if $X_t > S$ with the value of the option at time T being $F(X_t)$ where

$$F(x) = \max(x - S, 0) \tag{5.2}$$

We also suppose that there exists an inflation rate $r(t, x)$ such that R_0 dollars at time 0 is worth R_t at time t. Then,

$$dR_t = r(t, S_t) R_t dt$$

which implies

$$R_t = R_0 \exp\left\{ \int_0^t r(s, S_s)\, ds \right\}.$$

If $f(t, x)$ denotes the expected value of the payoff in time t dollars given $S_t = x$, then

$$f(t, x) = \mathbb{E}\left[\exp\left\{ -\int_t^T r(s, S_s)\, ds \right\} F(S_T) | S_t = x \right]. \tag{5.3}$$

Then assuming f is C^1 in t and C^2 in x, let us consider

$$M_t = \mathbb{E}[R_T^{-1} F(S_T) | \mathcal{F}_t].$$

Using the tower rule implies that if $s < t$,

$$\mathbb{E}[M_t | \mathcal{F}_t] = \mathbb{E}[\mathbb{E}(M_T | \mathcal{F}_t) | \mathcal{F}_s] = \mathbb{E}[M_T | \mathcal{F}_s] = M_s$$

which means that M_t is a martingale. We also know that R_t is $\mathcal{F}_t$-measurable and S_t is a Markov process, thus we can state

$$M_t = R_t^{-1} f(t, S_t) = R_t^{-1} \mathbb{E}\left[\exp\left\{ \int_t^T r(s, S_s)\, ds \right\} F(S_T) | S_t = x \right].$$

We then apply Ito's formula and note that R_t has finite variation to obtain

$$dM_t = d[R_t^{-1} f(t, x)] = f(t, x) d[R_t^{-1}] + R_t^{-1} df(t, x)$$

which implies the dt term of $d[R_t^{-1} f(t, x)]$ is

$$R_t^{-1}[-r(s, S_t) f(t, S_t) + \partial_t f(t, S_t) + m(t, S_t) \partial_x f(t, S_t) + \frac{1}{2} \sigma(t, S_t)^2 \partial_{xx} f(t, S_t)].$$

Since M_t is a martingale, by definition the dt term must be zero, which only occurs if

$$-r(s,S_t)f(t,S_t)+\partial_t f(t,S_t)+m(t,S_t)\partial_x f(t,S_t)+\frac{1}{2}\sigma(t,S_t)^2\partial_{xx}f(t,S_t)=0$$

Thus, we have obtained the following theorem.

Theorem 5.4. (Feynman-Kac Formula). *Suppose S_t is the price of a stock described by*

$$dS_t = m(t,S_t)dt+\sigma(t,S_t)dB_t$$

and $r(t,x)>0$ is the discounting rate. Suppose also that the payoff $F(S_T)$ at time T is given with $\mathbb{E}[|F(S_t)|]<\infty$. Then letting f be defined as in (5.3) with f C^1 in t and C^2 in x, it satisfies

$$\partial_t f(t,S_t)=-m(t,S_t)\partial_x f(t,S_t)-\frac{1}{2}\sigma(t,S_t)^2\partial_{xx}f(t,S_t)+r(s,S_t)f(t,S_t)$$

for $0\le t\ge T$ with end behavior f(T,x)=F(x)

The Feynman-Kac PDE is very similar to the Black-Scholes formula, however, it differs in its basic approach. Using Feynman-Kac, pricing an option according to its expected value in time t dollars can lead to arbitrage opportunities. This consequence should theoretically be impossible if the assets are priced correctly, and it leads to the intuition behind the Black-Scholes assumptions. The Black-Scholes model attempts to price options by creating a portfolio that creates zero arbitrage opportunities, implying a fair price. These features are further discussed in the next section where we derive the Black-Scholes formula.

6. Black-Scholes

Arbitrage is formally defined as the opportunity (positive probability) of making money with zero probability of losing money. This definition is the main tool used in setting up the Black-Scholes options pricing model. Consider a call option for a stock whose price moves according to a geometric Brownian motion.

Suppose the stock price S_t follows the geometric Brownian Motion,

$$dS_t = S_t[mdt+\sigma dB_t] \tag{6.1}$$

and also there exists a risk-free bound R_t satisfying

$$dR_t = rR_t dt \tag{6.2}$$

that is, $R_t = e^{rt}R_0$. Let T be a time in the future and suppose that we have the option to buy a share of stock at time T for strike price K. The value of this option at time T is

$$F(S_T)=\max(0,S_T-K).$$

The goal is to find the price $f(t,x)$ of the option at time $t<T$ given $S_t=x$.

One way of determining this price without using the Black-Scholes equation is to price the option purely by the expected profit at time t as shown below,

$$f(t,x)=\mathbb{E}[e^{-r(T-t)}F(S_T)|S_t=x].$$

In a previous section, we showed that this function satisfies the Feynman-Kac PDE,

$$\partial_t f(t,x)=rf(t,x)-mxf'(t,x)-\frac{\sigma^2x^2}{2}f''(t,x).$$

Using this model, however, if one sells an option at this price and buys a bond at the current interest rate, then there is a positive probability of losing money. This observation highlights a fundamental difference in the no arbitrage assumption of the Black-Scholes model.

The Black-Scholes approach to defining $f(t,x)$ is to let it be the value of a portfolio at time t given $S_t = x$ such that it can be hedged in order to guarantee a portfolio value $F(S_T)$ at time T. This portfolio consists of a ratio of stocks and bonds. Let V_t be the value of the portfolio at time t and a_t, b_t denote the ratio of stocks and bonds. Then,

$$V_t = a_t S_t + b_t R_t \tag{6.3}$$

In order to maintain the assumption of no arbitrage, we will manage the portfolio by switching the ratio of stocks and bonds such that no matter how the stock price moves, the value at time T will be

$$V_T = \max(0, S_T - K).$$

We also assume that the portfolio does not add outside resources, with the direct mathematical consequence that the change in the price of the portfolio is only dependent on the change in price of the assets,

$$dV_t = a_t dS_t + b_t dR_t. \tag{6.4}$$

If we assume (6.4) and plug in (6.1) and (6.1) then we get

$$\begin{aligned} dV_t &= a_t S_t[mdt + \sigma dB_t] + b_t r R_t dt \\ &= a_t S_t[mdt + \sigma dB_t] + r[V_t - a_t S_t]dt \\ &= [m a_t S_t + r(V_t - a_t S_t)]dt + \sigma a_t S_t dB_t. \end{aligned} \tag{6.5}$$

Alternatively, letting $V_t = f(S_t, t)$ as defined and assuming f is sufficiently differentiable, we can apply Ito's formula to show that

$$dV_t = df(t,S_t) = \partial_t f(t,S_t)dt + \partial_x f(t,S_t)dS_t + \frac{S_t^2\sigma^2}{2}\partial_{xx} f(t,S_t)dt$$

$$= [\partial_t f(t,S_t) + mS_t\partial_x f(t,S_t) + \frac{S_t^2\sigma^2}{2}\partial_{xx} f(t,S_t)]dt + \sigma S_t \partial_x f(t,S_t)dB_t. \tag{6.6}$$

Then by taking (6.5) and (6.6) and equating the dB_t terms, we see that the portfolio is defined by

$$a_t = \partial_x f(t,S_t), b_t = \frac{V_t - a_t S_t}{R_t}. \tag{6.7}$$

Equating the dt terms and plugging in (6.7), we get the Black-Scholes equation

$$\partial_t f(t,x) = rf(t,x) - rx\partial_x f(t,x) - \frac{\sigma^2 x^2}{2}\partial_{xx} f(t,x). \tag{6.8}$$

Note that the drift term m does not appear in this equation. Intuitively, this absence is because our price is based on the ability to hedge our portfolio to be an exact value at time T. Additionally, this gives the same formula as the Feynman-Kac formula except that m is replaced by r. This shows that when $m \neq r$ there exists arbitrage in the Feynman-Kac formula.

Using these facts, we can derive the Black-Scholes formula. We will not provide the full derivation since it is beyond the scope of this paper. The equation below shows the price of a European call option.

(6.9)

$$f(T-t,x) = x\phi\left(\frac{log(x/K)+(r+\frac{\sigma^2}{2})t}{\sigma\sqrt{t}}\right) - Ke^{-rt}\sigma\left(\frac{log(x/K)+(r-\frac{\sigma^2}{2})t}{\sigma\sqrt{t}}\right)$$

7. Heston Model

One of the weaknesses of the Black-Scholes model is that it assumes constant volatility of the option throughout its lifetime. In reality, the volatility of assets varies over time depending on a number of different factors. In short term pricing models, one of the most important factors is changing implied volatility. A modification to the constant volatility assumption of Black-Scholes is made in the Heston model. Instead of constant volatility, the volatility of the stock itself is a Brownian motion. It also differs from other stochastic volatility models in the following ways:

- It factors in possible correlation between the asset's price and its volatility.
- The volatility is mean reverting.
- It has an existing closed-form solution.
- It doesn't require the stock to follow a log-normal distribution.

The basic setup is as follows:

$$dS_t = rS_t dt + \sqrt{v_t} S_t dW_1 \tag{7.1}$$

$$dv_t = k(\theta - v_t)dt + \sigma\sqrt{v_t}dW_2$$

$$\mathbb{E}[dW_1, dW_2] = \rho dt$$

where S_t is the underlying asset price, σ is the volatility of $\sqrt{V_t}$, θ is the long term price variance, k is the rate reversion to θ, and W_1, W_2 are Brownian motions corresponding to the asset price and the asset's price volatility.

We use the same no arbitrage assumptions of Black-Scholes in attempting to create a portfolio that ensures a specific price in the future. We construct a portfolio consisting of one option $V = V(S, v, t)$, Δ units of the underlying stock S, and ϕ units of another option $U = U(S, v, t)$. The purpose of the second option is to hedge the volatility of the first option and stock. Then, from assuming a self financing portfolio similar to the Black-Scholes model, the portfolio value changes as follows:

$$d\Pi = dV + \delta dS + \phi dU.$$

We would now like to apply Ito's Lemma to dV. We differentiate with respect to t, S, v. Then, we get

$$dV = \frac{\partial V}{\partial t}dt + \frac{\partial V}{\partial S}dS + \frac{\partial V}{\partial v}dv + \frac{1}{2}vS^2\frac{\partial^2 V}{\partial S^2}dt + \frac{1}{2}\sigma^2 v\frac{\partial^2 V}{\partial v^2}dt + \sigma v\rho S\frac{\partial^2 V}{\partial v \partial S}dt.$$

If we apply Ito's Lemma to dU we get the same result, but in U. If we combine these two equations with (7.1), we can write

$$\begin{aligned}
d\Pi &= dV + \Delta dS + \phi dU \\
&= \left\{ \frac{\partial V}{\partial t} + \frac{1}{2} v S^2 \frac{\partial^2 V}{\partial S^2} + \rho\sigma v S \frac{\partial^2 V}{\partial v \partial S} + \frac{1}{2}\sigma^2 v \frac{\partial^2 V}{\partial^2} \right\} dt + \\
&\phi \left\{ \frac{\partial U}{\partial t} + \frac{1}{2} v S^2 \frac{\partial^2 U}{\partial S^2} + \rho\sigma v S \frac{\partial^2 U}{\partial v \partial S} + \frac{1}{2}\sigma^2 v \frac{\partial^2 U}{\partial^2} \right\} dt + \\
&\left\{ \frac{\partial V}{\partial S} + \phi \frac{\partial U}{\partial S} + \Delta \right\} dS + \left\{ \frac{\partial V}{\partial v} + \phi \frac{\partial U}{\partial v} \right\} dv. \qquad (7.2)
\end{aligned}$$

In order to meet our no arbitrage requirement, the portfolio must be hedged against the stock price and volatility, meaning that the last two terms of (7.2) with coefficients dS, dv must be zero. This observation gives us the following hedge parameters:

$$\phi = -\frac{\frac{\partial V}{\partial v}}{\frac{\partial U}{\partial v}}$$

$$\Delta = -\phi \frac{\partial U}{\partial S} - \frac{\partial V}{\partial S}. \qquad (7.3)$$

We also assume that the portfolio must earn the risk free rate r, thus, $d\Pi = r\Pi dt$. With the values and assumptions used to derive (7.3), we get

$$\begin{aligned}
d\Pi = &\left\{ \frac{\partial V}{\partial t} + \frac{1}{2} v S^2 \frac{\partial^2 V}{\partial S^2} + \rho\sigma v S \frac{\partial^2 V}{\partial v \partial S} + \frac{1}{2}\sigma^2 v \frac{\partial^2 V}{\partial^2} \right\} dt + \\
&\phi \left\{ \frac{\partial U}{\partial t} + \frac{1}{2} v S^2 \frac{\partial^2 U}{\partial S^2} + \rho\sigma v S \frac{\partial^2 U}{\partial v \partial S} + \frac{1}{2}\sigma^2 v \frac{\partial^2 U}{\partial v^2} \right\} dt
\end{aligned}$$

which we can write as $d\Pi = (A + \phi B)dt$. Then, we get

$$A + \phi B = r(V + \Delta S + \phi U).$$

Substituting in the portfolio values, we get

$$\frac{A - rV + rS\frac{\partial V}{\partial S}}{\frac{\partial V}{\partial v}} = \frac{B - rU + rS\frac{\partial U}{\partial S}}{\frac{\partial U}{\partial v}}. \qquad (7.4)$$

Looking at (7.4), we notice the left hand side is only a function of V and the right side is only a function of U. This means that both sides can be written as some function $f(S, v, t)$. Using what we specified earlier, let $f(S, v, t) = -k(\theta - v) + \lambda(S, v, t)$ where $\lambda(S, v, t)$ is the volatility risk price. If we substitute f into the left side of (7.4) and substitute for B, we arrive at the Heston PDE expressed in terms of the underlying asset price S

$$\begin{aligned}
&\frac{\partial U}{\partial t} + \frac{1}{2} v S^2 \frac{\partial^2 U}{\partial S^2} + \rho\sigma v S \frac{\partial^2 U}{\partial v \partial S} + \frac{1}{2}\sigma^2 v \frac{\partial^2 U}{\partial v^2} \\
&-rU + rS\frac{\partial U}{\partial S} + [k(\theta - v) - \lambda(S, v, t)]\frac{\partial U}{\partial v} = 0 \qquad (7.5)
\end{aligned}$$

This form of the PDE is quite complicated to use so we derive the PDE in terms of the log price of the asset.

Let $x = \ln S$. Then we want to express the PDE in terms of x, t, v. We require the following derivatives, which can be derived without too much difficulty

$$\frac{\partial U}{\partial S}, \frac{\partial^2 U}{\partial v \partial S}, \frac{\partial^2 U}{\partial S^2}.$$

We plug these values into (7.5) and get

$$\frac{\partial U}{\partial t} + \frac{1}{2}v\frac{\partial^2 U}{\partial x^2} + \left(r - \frac{1}{2}v\right)\frac{\partial U}{\partial x} + \rho\sigma v\frac{\partial^2 U}{\partial v \partial x} +$$
$$\frac{1}{2}\sigma^2 v\frac{\partial^2 U}{\partial v^2} - rU + rS\frac{\partial U}{\partial S} + [k(\theta - v) - \lambda v]\frac{\partial U}{\partial v} = 0.$$

In the Heston model, we set the market price of risk to be a linear function of volatility so we let $\lambda(S, v, t) = \lambda v$. It is possible to derive a closed form for the price of a call option. However it gives a complex analytic answer that doesn't provide too much intuition and is mainly important if you want to actually calculate the price. The solution ends up being closed, but requires a numerical method to compute the integral. As you can see, this form of the Heston PDE contains every term in the Black-Scholes formula, but it also has extra terms to deal with the additional option in the portfolio.

From an analytical perspective, the Heston model has been shown to overall outperform the Black-Scholes model in options pricing. Specifically, the Heston model is able to better capture the price of long term options where volatility may change in a specific manner. In terms of actual use, the Heston model requires many parameters that are tuned by the current market. Thus, if one were to use Black-Scholes options prices to calibrate the Heston model, then they would give the same price. Conversely, if you had the Heston parameters and price, you could find the volatility and plug it into Black-Scholes to find the same price. Thus, much of the utility of the Heston model arises from further tuning its inputs in the current market, giving it a greater degree of control than the Black-Scholes model.

Acknowledgments

It is a pleasure to thank my mentor Devon Ingram for his insightful guidance over the course of writing this paper along with providing many of the materials required to write it.

References

[1] Gregory F. Lawler. Stochastic Calculus: An Introduction with Applications. http://www.math.uchicago.edu/ lawler/finbook2.pdf.
[2] Gregory F. Lawler. Notes On Probability. http://www.math.uchicago.edu/ lawler/probnotes.pdf.
[3] Fischer Black and Myron Scholes. The Pricing of Options and Corporate Liabilities. https://www.cs.princeton.edu/courses/archive/fall09/cos323/papers/black scholes73.pdf.
[4] Brandon Washburn and Mehmet Dik. Derivation of Black-Scholes Equation using Ito's Lemma. https://doi.org/10.47086/pims.956201.
[5] Fabrice Douglas Rouah. Simplified Derivation of the Heston Model. https://frouah.com/finance
[6] Robin Dunn. Estimating Option Prices with Heston's Stochastic Volatility Model. https://www.valpo.edu/mathematics-statistics/files/2015/07/Estimating-Option-Prices-with-Heston
[7] John C. Hull. Options, Futures, and Other Derivatives. Pearson. 1989