TITLE: Portfolio Optimization
- A Stochastic Control Approach

AUTHOR: Aaron Brooks-Moorcroft

SUPERVISOR: Dr. Gennady Shaikhet

DATE: May 1, 2019
## Contents

1 Preliminaries  
1.1 Stochastic Process ........................................... 6  
1.2 sigma-field .................................................. 6  
1.3 Filtration .................................................... 7  
1.4 Brownian Motion ............................................. 8  
1.5 Brownian Motion with Drift ................................. 9  
1.6 Geometric Brownian Motion ................................. 9  
1.7 Martingales .................................................. 10  

2 Simulations: 1 .................................................. 13  

3 Stock Model derivation  
3.1 Characteristics .............................................. 16  
3.2 Utility Theory ............................................... 18  
3.2.1 Assumed Utility function ............................... 19  

4 Elements of Stochastic Calculus  
4.1 Defining the Stochastic Integral ............................ 20  
4.1.1 Stochastic Integral is a Martingale .................... 22  
4.2 Itô’s lemma .................................................. 23  
4.2.1 Example: SDE representation of GBM process ...... 26  

5 Stochastic Control  
5.1 Derivation of the HJB Equation ............................ 28  
5.2 The value function ......................................... 30  
5.3 Optimization over \( \theta \) .................................. 32  
5.4 Optimization over \( c \) ..................................... 34  
5.5 Optimization over \( \theta \) and \( c \) ......................... 35  
5.6 Formal statement of the problem ......................... 36  

6 Simulations: 2 .................................................. 36
7 Simulations: 3 - A brief study on five Technology Giants

Appendices

A MATLAB Code .......................................................... 44
Acknowledgements

Only with the tremendous knowledge, guidance, and support provided to me during my studies at Carleton University could I write this text. I thank every member who has made an impact on me along the way. I would especially like to thank Dr. Shaikhet for his motivation and incredibly detailed, yet accessible explanations on topics which could take several lifetimes to fully understand.

Lastly, my family and friends for supporting me through my academic journey.
"Risk comes from not knowing what you’re doing."

WARREN BUFFETT
Abstract. This text aims to provide the reader with a summary of key concepts surrounding the classical Merton problem. It begins by introducing concepts of Brownian motion and martingales and moves on to expressing the dynamics of an investor’s wealth equation subject to a cost function. Next, we digress into stochastic calculus which provides us with the tools to express stock processes. Finally, we apply these concepts which will allow for the construction of the Hamilton-Jacobi-Bellman equation for the value function which we will try to optimize.

1 Preliminaries

Before entering the world of stochastic control processes which allow us to define optimal allocations of wealth based on certain constraints, it is necessary for us to define some common concepts which will be found throughout this text. We will begin by defining the overall landscape of stochastic processes and how they are defined and interpreted. We will move on to discuss Brownian Motion and how it applies to financial mathematics along with some useful variants of the process. After that, we will have the essentials which will be carried forward into more complex analysis and optimization problems.

1.1 Stochastic Process

Stochastic processes are fundamental in the creation of mathematical models for financial markets. They are useful since they model an event at each point in time starting at some time $t$ similar to deterministic processes. A stochastic process can be stated as a random function $X$ so that:

$$X = \{X_t; t \geq 0\} \equiv \{X(t); t \geq 0\}$$ (1)

The reason we use stochastic models will be realized in subsequent sections.

1.2 $\sigma$ - field

Definition 1.1 ($\sigma$ - field). A $\sigma$ - field $\mathcal{F}$ is a collection of subsets of $X$ which satisfy the conditions outlined below:
i. \( \emptyset \in \mathcal{F} \)

ii. \( \mathcal{F} \) is closed under complementation. That is, if \( A \in \mathcal{F} \Rightarrow A^C \in \mathcal{F} \)

iii. \( \mathcal{F} \) is closed under countable unions. That is if \( A_1, A_2, A_3, \ldots \) are in \( \mathcal{F} \) then so is \( \bigcup_{i=1}^{\infty} A_i \)

We regard the \( \sigma \)-field generated by stochastic process \( X_t \) as \( \sigma(X_t) \) to be the smallest \( \sigma \)-field which contains all information that can be obtained through paths of \( X \) over the interval \( [0,t] \).

### 1.3 Filtration

Given a generalized idea that historical asset prices are used to determine future ones, it is evident that a filtration is an important concept.

**Definition 1.2 (Filtration).** Let \( \mathcal{F}_T = (\mathcal{F}_t)_{t \in [0,T]} \) be an increasing set of \( \sigma \)-fields. Then, if \( \forall 0 \leq s \leq t \):

\[
\mathcal{F}_s \subseteq \mathcal{F}_t
\]

Then \( \mathcal{F}_T \) is called a filtration.

We interpret each \( \mathcal{F}_t \) as the information known at time \( t \). When we consider some time \( s \leq t \) we would have greater than or equal to the amount of information known at later time \( t \) than that of \( s \). An example of this information can be regarded as the history of a stock price over time. As time progresses you would have more information about the price than you would have the previous day. The connection between stochastic processes and filtrations is obvious when we consider a set \( \{(X_t, \mathcal{F}_t)\}, t \geq 0 \in S \) where \( X_t \in \mathbb{R} \) is a random variable which is \( \mathcal{F} \)-measurable and \( \mathcal{F}_t \) is a filtration. Such a set is called a stochastic process with filtration.

**Definition 1.3 (Adapted process).** A stochastic process \( X_t \) is said to have adapted to filtration \( (\mathcal{F}_t, t \geq 0) \) if:

\[
\sigma(X_t) \subseteq \mathcal{F}_t, \ \forall t \geq 0
\]

That is, \( X_t \) is an adapted process.
Remark. The stochastic process $X_t$ must be adapted to the natural filtration generated by $X$ so that:

$$\mathcal{F}_t = \sigma(X_t, s \leq t)$$

which shows that $X$ does not contain more information than $\mathcal{F}_t$.

The principle of filtration will be largely used in martingale processes which will be discussed in later sections of this text.

1.4 Brownian Motion

Brownian Motion (BM) is a term coined by botanist Robert Brown in 1827 to model the irregular motion of pollen particles suspended in fluid. This model in conjunction with the heat equation has been adapted since as early as the year 1900 and is still largely used today to model noise in financial markets. Today, many of the key results in financial modelling rely on concepts of Brownian Motion to model stock prices.

Hence, we must begin by defining Brownian Motion for one dimension as follows:

**Definition 1.4 (One-dimensional Brownian Motion).** Let $W$ be a stochastic process $W = \{W(t); t \geq 0\}$. Then $W$ is said to be a BM process if:

i. $W(0) = 0$

ii. $\{W(t)\}$ has stationary increments:

$$W(t) - W(s) \sim N(0, t - s), \ 0 \leq s \leq t$$

iii. $\{W(t)\}$ has independent increments:

$$W(d) - W(c) \text{ is independent of } W(b) - W(a), \ 0 \leq a \leq b \leq c \leq d$$

Lastly, if we define a fourth condition such that:

iv. $W(t) \sim N(0, \sigma^2 t), \ \forall t > 0$
This BM process is called a Wiener process. Lastly, a standard BM process is obtained upon setting \( \sigma^2 = 1 \). When \( \sigma^2 \neq 1 \) we can achieve a standard BM process by setting \( Y(t) = X(t)/\sigma \)

**Definition 1.5** (n-dimensional Brownian Motion). Taking an \( n \)-dimensional set \( \{X(t)_n\} \), \( n \geq 0 \) where each \( X(t)_i \), \( 0 \leq i \leq n \) is a one-dimensional BM process, then this would be called an \( n \)-dimensional BM process.

An implication from the definition above is that each coordinate of \( X(t)_i \) will be a standard BM and independent.

### 1.5 Brownian Motion with Drift

**Definition 1.6** (Brownian Motion with drift). Let \( X \) be a BM process \( \{X(t); t \geq 0\} \). Then \( X \) is said to be a BM process with drift coefficient \( \mu \) and volatility \( \sigma \) expressed as:

\[
X_t := \mu t + \sigma W_t
\]  

Here we note that \( X(t) \sim N(\mu t, \sigma^2 t) \) as the mean now varies over time.

**Remark.** Once again, taking an \( n \)-dimensional set \( \{X(t)_n\} \), \( n > 0 \) where each \( X(t)_i \), , \( 0 < i \leq n \) is a one-dimensional BM process with drift, then this known as an \( n \)-dimensional BM process with drift which is also a generalized Wiener process. Similarly, as with the case of Brownian Motion without drift, each coordinate of \( X(t)_i \) will be a standard BM with drift and independent.

### 1.6 Geometric Brownian Motion

Another important concept in Brownian motion is Geometric Brownian motion (GBM). This concept is often used to model stock prices over time and is also useful since the process only has non-negative values similar to stock prices. An example of this would be the Black-Scholes formula for pricing options. Below is the formal definition of a GBM process.

**Definition 1.7** (Geometric Brownian Motion). Let \( \{W(t), t \geq 0\} \) be a standard Brownian motion process and \( \mu \) and \( \sigma^2 \) be the drift and variance coefficients respectively. Then we say
that the process \( \{S(t), t \geq 0\} \) is a *Geometric Brownian motion* process:

\[
S(t) = S_0 e^{(\mu - \sigma^2/2)t + \sigma W_t}
\] (3)

There are many useful properties of a GBM process that can be derived. One can condition on the GBM equation given its' history at time \( 0 < s < t \) and apply its' normally distributed properties for process \( S(t) \). These will be derived in subsequent sections after Martingales are defined.

### 1.7 Martingales

A useful concept related to the generalized Wiener process is that of Martingales. Martingales are often used for games of chance. Consider a casino game of Blackjack in which the odds of an individual winning when playing against the house is set at around 48%. Then the expected long-term outcome would be lower than that of the initial value of the individual’s wealth. On the other hand, the house stands a 52% chance of winning and therefore will have an expected wealth greater than that of their original wealth. The above ideas are highly applicable to those in the stock market as prices are thought to have upward drift.

Below is a formal definition a Martingale and different scenarios with respect to the drift coefficient \( \mu \).

**Definition 1.8** (Martingale). Let \( X(t) \) be a process \( \{(X_t, \mathcal{F}_t)\}, t > 0 \in S \) where \( X_t \in \mathbb{R} \) and \( S \) is an ordered set. Then:

i \( \mathbb{E}(X_t | \mathcal{F}_s) = X_s \)

Then: \( \{(X_t, \mathcal{F}_t)\}, \) is a *martingale*

ii \( \mathbb{E}(X_t | \mathcal{F}_s) \leq X_s \)

Then: \( \{(X_t, \mathcal{F}_t)\}, \) is a *super-martingale*

iii \( \mathbb{E}(X_t | \mathcal{F}_s) \geq X_s \)

Then: \( \{(X_t, \mathcal{F}_t)\}, \) is a *sub-martingale*
Back to the example above, the individual playing against the casino would see themselves having a super-martingale process as the expected worth at the end of playing would be less than that having not played. The casino would have a sub-martingale process as their expected wealth increases with each iteration of the game. Finally, consider a game with 50/50 odds such as the number being even versus odd when rolling a die. If the individual wins one dollar when an even number is rolled and pays one dollar when an odd number rolled then such an individual would be considered to follow a martingale process.

Hence, in our definition of one-dimensional Brownian Motion $W_t$ we see for $s \leq t$ that:

$$\mathbb{E}(W_t | F_s) = \mathbb{E}(W_t + (W_s - W_s) | F_s)$$

$$= \mathbb{E}(W_t - W_s | F_s) + W_s$$

and since the increment $W_t - W_s$ is independent of $F_s$

$$= \mathbb{E}(W_t - W_s) + W_s$$

$$= W_s$$

Hence, the one-dimensional Brownian Motion $W_t$ is a martingale.

When considering an equation such as $X_t = \mu t + \sigma W_t$, $\mu, \sigma \in \mathbb{R}$ it is easy to see that when $\mu = 0 \Rightarrow X_t$ is a martingale process, when $\mu > 0 \Rightarrow X_t$ is a sub-martingale process, and when $\mu < 0 \Rightarrow X_t$ is a super-martingale process.

Revisiting the concept of Geometric Brownian motion, we consider a GBM process $S_t$. One can condition on the GBM equation given its history at time $0 < s < t$ and apply its normally distributed properties to derive useful results.

Let $S_t, t \geq 0$ be a GBM process with filtration $\mathcal{F}_t$. Then define history of the process up until time $s$ as $(S_k, 0 \leq k \leq s)$ with filtration $\mathcal{F}_s$ and $\{X_t, t \geq 0\}$ a BM process with drift coefficient $\mu$ and variance $\sigma^2$. Now:

$$\mathbb{E}[S_t | \mathcal{F}_s, 0 \leq k \leq s] = \mathbb{E}[e^{X_t} | X_k, 0 \leq k \leq s]$$

$$= e^{X_s} \mathbb{E}[e^{X_t-X_s} | X_k, 0 \leq k \leq s]$$

$$= e^{X_s} \mathbb{E}[e^{X_t-X_s}]$$

Since $X_t - X_s \sim N(\mu(t-s), \sigma^2(t-s))$
(and by the stationary and independent increments of BM with drift):

$$\Rightarrow \mathbb{E}[S_t | \mathcal{F}_s, 0 \leq k \leq s] = S_s e^{(t-s)(\mu + \sigma^2/2)}$$

(4)

The result from (4) will be useful in subsequent sections.
2 Simulations: 1

In this section we simulate Brownian Motion and Geometric Brownian Motion and observe the results.

We begin with a basic simulation of Brownian Motion over the interval $[0, 1]$. The algorithm is as follows:

![Algorithm 1: Stochastic Brownian motion process $X = \{X_t, 0 \leq t \leq 1\}$](Algorithm1.png)

1. **Set Initial Conditions**: Set $n$-observations large (such as $n=10000$);
2. **Generate**: vector $A = (Z_1, Z_2, \ldots, Z_n)$ where each $Z_i \sim N(0, 1)$;
3. **for** $i = 1$ **to** $n$ **do**
   4. $X(0) = 0$;
   5. $X(t_1) = \sqrt{\Delta}Z_1$;
   6. $X(t_2) = \sqrt{\Delta}Z_2 + w(t_1)$;
   7. $X(t_n) = \sqrt{\Delta}Z_n + w(t_{n-1})$;
4. **end**
5. **Plot**: $[0 : 1/N : 1], [0, X_t]$

Figure 1: Simulated path of Brownian Motion using MATLAB
Figure 1 shows that the simulation of the path \( X(t) \) demonstrates all of the properties of a standard Brownian Motion process.

Next, we look to simulate the effects that different values of coefficient \( \mu \) has on Brownian Motion with drift. The process for simulating such processes is as follows:

<table>
<thead>
<tr>
<th>Algorithm 2: Brownian motion process with drift ( X = { X_t, 0 \leq t \leq 3 } )</th>
</tr>
</thead>
</table>
| 1 Set Initial Conditions: \( T = 3 \) n-large (365·T) \( \mu = [0.5, 0, -0.5] \) \( \sigma = [0.25, 1] \)  
\( S_0 = 100 \) \( dt = T/n \) \( d = 1000 \); |
| 2 for \( j = 1 \) to size(\( \sigma \)) do |
| 3 \( \sigma' = \sigma(j); \) |
| 4 for \( k = 1 \) to size(\( \mu \)) do |
| 5 \( \mu' = \mu(k); \) |
| 6 \( S^d = S_0 \) (a \( d \)-dimensional row vector of initial stock price); |
| 7 for \( t = 1 \) to \( n \) do |
| 8 for \( i = 1 \) to \( d \) do |
| Update stock price at every step according to GBM process; |
| Generate \( W \), random standard normal r.v. |
| \( S^d_{t+1} = S^d_t \cdot e^{(\mu' - 0.5 \cdot \sigma'^2) \cdot dt + \sigma' \cdot \sqrt{dt} \cdot W}; \) |
| end |
| end |
| Compute: Mean path at each step, \( M' \); |
| Plot: \([0 : dt : T], [0, M_t] \) |
| end |

Figure 2 demonstrates the extent to which the effects caused by the drift parameter \( \mu \) has on a stock model. Clearly, we see here how a positive \( \mu \) will create a sub-martingale process. Conversely, a negative \( \mu \) will create a super-martingale process. Lastly, by observing Figure 3 even though an average path for over 1000 stock prices were simulated, the volatility, \( \sigma \), had a significant effect at some points in time. Consider an exchange traded fund (ETF) which tracks a major index such at the S&P 500. On average we tend to see the value
Figure 2: Simulated average of 1000 stock prices for different values of $\mu$, $\sigma = 0.25$

increase over time as stocks tend to drift upward ($\mu > 0$). However at times we may see increased volatility such as major news events.

Figure 3: Simulated average of 1000 stock prices for different values of $\mu$, $\sigma = 1$
3 Stock Model derivation

In this next section, the foundations for creating a stochastic stock model will be defined. Key concepts in financial mathematics will be derived that will ultimately lead to the goal of solving what is known as the Hamilton Jacobi-Bellman (HJB) equation. There are many factors to account for, transaction costs, bankruptcy costs, and investor habit just to name a few. However, all these complications will be set aside for simplicity as computations can become quite intense from the get-go.

We will begin by defining the characteristics of an agent’s wealth and move on to Utility Theory which explains how an agent allocates a scarce resource when there is risk involved. Next we will discuss the stochastic integral and Itô’s lemma which are at the heart of stochastic control problems.

3.1 Characteristics

We will now begin by defining some variables which will be used in the wealth equation. The wealth equation is used to describe the dynamics of their wealth over time. So, at time \( t \) we consider an investor to have:

- \( w_t \): wealth of portfolio
- \( S_t \): asset price
- \( n_t \): portfolio process (number of assets)
- \( \delta_t \): dividends\(^1\)
- \( e_t \): scalar endowment/rejuvenation process\(^2\)
- \( c_t \): scalar consumption
- \( r_t \): risk-free interest rate

\(^1\)Often omitted for simplicity
\(^2\)Often omitted for simplicity
Consider an individual’s wealth at time zero equal to $w_0$. Then at time $t$ one can express the dynamics of their wealth in an expression such as:

$$dw_t = r_t w_t dt + n_t dS_t - n_t r_t S_t dt + n_t \delta dt + e_t dt - c_t dt$$

(5)

where the differentials represent increments of the wealth process over an infinitesimal time increment $dt$:

$$dw_t = w_{t+dt} - w_t$$

collecting terms in (5):

$$= r_t (w_t - n_t \cdot S_t) dt + n_t \cdot (dS_t + \delta dt) + e_t dt - c_t dt$$

(6)

In this text however, for simplicity we will define the wealth equation with fewer terms. Also note that we express the dynamics in terms of $X_t$ rather than $w_t$.

**Definition 3.1** (Wealth equation). Let $X_t$ be the wealth of a portfolio at time $t$ when starting with wealth at time zero $X_0 = x$. Then the dynamics of an investor’s wealth can be expressed by the process where $S_t$ is modelled using Geometric Brownian Motion is given by:

$$dX_t = r_t X_t dt + n_t dS_t - n_t r_t S_t dt - c_t dt$$

(7)

That is, we omit any endowment and dividends from the process. From this point onwards we assume that stock prices follow Geometric Brownian Motion processes. Similarly, where $dX_t = X_{t+dt} - X_t$ for small incremental time.

Generally, one would assume that the initial wealth $w_0$ is known as with the processes $\delta, S, r, e$ since they can either be found at any given moment in the open market or are known to the investor. One can easily interpret this equation in the following way. Suppose that an investor decides to invest in zero risky\(^3\) assets. Then $n = 0$ so that an investor will earn $dw_t = r_t w_t dt + e_t dt - c_t dt$. This function shows that the investor’s change in wealth is a function of their wealth multiplied by the risk-free rate plus some endowment.

---

\(^3\)The term ‘risky’ is not used in the sense of a measure of risk for a stock but rather investing in assets other than simply earning the ‘risk-free’ rate
amount less their consumption rate. So to speak, this represents the average person who does not invest keeping all of their cash in their bank account paying their bills, buying their groceries and depositing their cheques. Consider now an individual deciding to invest in $n$ risky assets at time zero. Then, at time $t$ the individual's wealth would be comprised of $\sum_{i=1}^{d} n_i^t S_i^t$, $i = 1, \ldots, d$ consists of $d$ investments of $n_i^t$ holdings at market value $S_i^t$ plus dividends received in the amount of $\sum_{i=1}^{d} n_i^0 \delta_i^t$. In addition, the individual would earn the risk-free rate on the cash holdings equal to $w_0 - \sum_{i=1}^{d} n_i^0 S_i^t$ plus some endowment amount less their consumption. The concept of the risk-free rate is theoretical since there does not exist an investment with zero-risk. However there exist proxies such as US three-month Treasury bills which are government backed. Hence, T-bills can be used to proxy the risk-free rate and are used to describe the minimum return on investment that an investor should require. It is often the convention to define some identities defining the worth of combined assets along with the proportion of wealth for any given asset $i$. So we define:

At time $t$ asset $i$ has wealth $\theta_i^t$ so that:

$$\theta_i^t = n_i^t S_i^t$$  \hfill (8)

Also, at time $t$ asset $i$ represents proportion of wealth $\pi_i^t$ so that:

$$\theta_i^t = \pi_i^t w_t$$  \hfill (9)

Where $\sum_{i=1}^{d} \pi_i^t w_t = 1$ when all of an investor's funds are invested in risky assets.

### 3.2 Utility Theory

An investor's goal is to maximize their returns. However, investors are rational and only tolerate a certain level of risk. Hence, an investor would like to maximize their returns subject to a tolerable level of risk. This idea can be summarized by a function called the utility function where each investor's utility function is unique. The original purpose of a utility function was to explain how an investor would prefer more of something to less, however, a loss of value $x$ would cause more hurt than the pleasure caused by a gain of $x$. Also, the more an individual has, the less they prefer.
An investor’s utility function is a concave increasing function. Since an investor prefers less when they have more, the function expresses diminishing returns.

An investor is said to have a utility function $u(x)$ (or $u(t,x), U$) which contain the following properties

\begin{enumerate}
  \item $x \succ y \iff \mathbb{E}(u(x)) > \mathbb{E}(u(y))$
  \item $u(\mathbb{E}(x)) \geq \mathbb{E}(u(x))$
\end{enumerate}

where $\succ$ represents an investor’s preference. Component (i) represents that an investor would prefer investment decision $x$ to alternative $y$ provided that investment $x$ provides superior returns to $y$. Component (ii) demonstrates that an investor’s expected utility is less than their utility of expected return. This satisfies Jensen’s inequality which shows the concavity of $u(x)$. Lastly, the implication that $u'(x) > 0$ shows $u(x)$ increasing and $u''(x)$ shows the decreasing marginal returns.

### 3.2.1 Assumed Utility function

The ultimate goal of the utility function is to provide a reasonable constraint to the optimal control problem. For this reason we will consider a utility function such as:

$$u(t, x) = e^{-\rho t} u(x)$$

where $u(x) = \frac{x^{1-R}}{1-R}$ and $R$ is an agent’s risk adversity parameter. An agent with neutral risk would have $R = 0$ and as $R \nearrow \infty$ would imply infinite risk adversity. This utility function is known as the constant-relative-risk-aversion (CRRA) form. We will assume this form as it will provide a constraint and eventually allow for us to arrive at a meaningful result without adding unnecessary complications.

Using the above results, we can now define an equation to represent the optimization problem.

An investor is driven to find a value function $V(w)$ which will maximize the expected value of their utility function. That is,

$$V(w_0) = \sup_{n,c} \mathbb{E} \left[ \int_{0}^{\infty} e^{-\rho t} \frac{c_t^{1-R}}{1-R} dt \right]$$

19
Remark. Here, it is important to note that we only consider reasonable pairs of \((n_t, c_t)\) such that satisfy an initial wealth amount of \(w_0\). In subsequent sections we will aim to solve simpler versions of (11) with the assumption that the stock process \(S_t\) follows a Geometric Brownian Motion process. Now, we will dive into stochastic control to gain the tools to do so.

4 Elements of Stochastic Calculus

Stochastic calculus is a fundamental topic in financial mathematics. It employs modern statistical techniques to develop models for stock prices. Deriving what is known as the stochastic integral uses many techniques mentioned previously to further our goal of constructing an HJB and solving for optimal control. We can also derive some useful methods to deal with the wealth dynamics described in (7). Next, we will impose some optimizing constraints which will allow the investor to achieve an optimum level of wealth to satisfy their own objective.

4.1 Defining the Stochastic Integral

Prior to defining the stochastic integral, it is worth investigating some more properties of Brownian motion which gives motivation to the result. We will examine ways to evaluate an integral containing a component of Brownian Motion \(W_t\), but will quickly realize that the standard techniques will be inadequate due to the nature of Brownian Motion.

First we note that an integral of some differentiable function \(F\) and process \(X(s)\) on \((\mathbb{R}, \mathbb{R})\) can be be computed for \(t > 0\) as:

\[
\int_0^t X_s dF_s = \int_0^t X_s \frac{dF_s}{ds} ds = \int_0^t X_s f_s ds
\]

However, this relies on the fact that \(F\) is differentiable. Unfortunately, Brownian motion \(W_t\) is nowhere differentiable as the path is not smooth and there are infinity many tangent points at every point. At the same time, we can try to also compute the integral for continuous \(X\)
using a Riemann-Stieltjes representation as:

\[
\int_0^t X_s dF_s = \lim_{n \to \infty} \sum_{i=1}^n X_{it/n} \cdot (F_{(i-1)t/n} - F_{it/n})
\]

However, this is only permissible when \( F \) has bounded variation and unfortunately, \( W \) has unbounded first variation, so we must develop another method to handle this.

**Definition 4.1** (Bounded Variation). A real-valued function \( f \) over the interval \([0, 1]\) is said to have \( p \)-bounded variation for some \( p > 0 \) when:

\[
\sup \sum_{i=1}^n |f(t_i) - f(t_{i-1})|^p < \infty,
\]

where the supremum is taken over each partition \( \tau \) of \([0, 1]\).

This result leads to questioning whether the definition of a stochastic integral given by:

\[
I_t(X) := \int_0^t X_s dW_s
\]

can be generalized when we integrate over one-dimensional Brownian Motion since directly applying the integration techniques above will be erroneous as Brownian motion is almost-surely nowhere differentiable. This is also caused by a variation of infinite size over any interval \([s, t], 0 \leq s < t < \infty\). Hence, we must define the stochastic integral for a certain process \( X_t \); such that \( X_t \) must be a *simple process*. These so-called simple integrals are a great tool to understand the properties of the general stochastic integrals.

**Definition 4.2** (Simple Stochastic Process). A simple stochastic process is a stochastic process \( X_t \) where \( 0 = t_0 < t_1 < \ldots < t_n, n \in \mathbb{N} \) is a partition over \([0, T]\) so that \( X_t(w), \omega \in \Omega \) is represented by:

\[
X_t = X_t(w) = X(t, w) = \Delta_0(w)1_0(t) + \sum_{i=1}^n \Delta_i(w)1_{(t_{i-1}, t_i)}(t)
\]

**Remark.** An important result from the above definition is that \( X_t \) is a left-continuous step function. The height of each step is \( \Delta_i(w)1_{(t_{i-1}, t_i)}(t) \). Figure 4 is a graphical representation for a simulated simple stochastic process \( X_t \). A stochastic integral can be defined with respect to some stochastic process\(^4\) \( X \) which is a limit of Riemann sequences.

\(^4\)Used interchangeably as a simple process
Therefore, we can finally express the stochastic integral with which contains Brownian Motion.

**Definition 4.3 (Stochastic Integral).** Let \( \{X_t\}_{t \in [0, T]} \) be a simple process. Then the Stochastic Integral \( I_t(X) \) for \( t \in (t_k, t_{k+1}] \) is given as:

\[
I_t(X) := \int_0^t X_s dW_s = \sum_{i=0}^{k-1} \Delta(t_i)[W(t_{i+1}) - W(t_i)] + \Delta(s)[W(t) - W(s)]
\]  

(14)

where \( W_t - W_s \) is some incremental Brownian motion for \( s < t \) and \( X \) is a constant.

### 4.1.1 Stochastic Integral is a Martingale

Consider the simple Process \( \{X_t\}_{t \in [0, T]} \). Then the stochastic integral can be written for \( 0 \leq s \leq t \) as:

\[
I(t) = \sum_{i=0}^{k-1} \Delta(t_i)[W(t_{i+1}) - W(t_i)] + \Delta(s) \underbrace{[W(t) - W(s)]}_{I(s)} + \underbrace{\text{incrementalBM}}_{t}
\]
taking conditional expectation given filtration $\mathcal{F}$:

$$
\mathbb{E}[I(t)|\mathcal{F}] = I(s)
\Rightarrow I(t) = I(s) + \Delta(s)[W(t) - W(s)]
$$

simply showing $a$ has expected value zero so that:

$$
\mathbb{E}[I(t)|\mathcal{F}] = \mathbb{E}[I(s) + \Delta(s)(W(t) - W(s))|\mathcal{F}]
= I(s) + \mathbb{E}[\Delta(s)(W(t) - W(s))|\mathcal{F}]
= I(s) + \Delta(s)\mathbb{E}[(W(t) - W(s))|\mathcal{F}]
= I(s)
$$

Which proves the result that the stochastic integral is a martingale. This of course assumes W.L.O.G that $t > 1$ and $s > 0$ however some simple modifications will also yield the same result. The fact that the stochastic integral is a martingale process shows that given the history for any stock one would assume that the price would remain the same at any time $t > s$.

### 4.2 Itô’s lemma

Now that we have the definition of the stochastic integral we can now define the general diffusion process. So, we consider the results derived in (14) in the definition below.

**Definition 4.4 (Itô Process).** Let $W$ be a Brownian Motion process. Then $X$ is an Itô process described by:

$$
X_t = X_0 + \int_0^t \mu(s, X_s)ds + \int_0^t \sigma(s, X_s)dW_s
$$

(15)
Or more compactly,
\[ S_t = S_0 + \int_0^t aS_s ds + \int_0^t \sigma S_s dW_s \] (16)

Where \( W_s \) is some form of Brownian Motion and \( a(t, x) \) and \( b(t, x) \) are deterministic functions\(^5\).

Formula (16), (15) are used to formulate a standard model for stock prices over the interval \( 0 \leq t \leq T \). There are two components to this model. The first integral on the right hand side of (15) is a Riemann integral and the second integral is an Itô stochastic integral described in (14). This equation is said to have a driving process by Brownian motion process \( W_t \).

Equation (15) often is expressed in differential form:
\[ dX_t = \mu(t, X_t) dt + \sigma(t, X_t) dW_t \] (17)

Where \( dx(t) \) and \( W_t \) represent Brownian motion. One can interpret these equations as for any change of \( dX_t = X_{t+dt} - X_t \) is a result of changes in time \( t \) by a factor of \( a(t, X_t) \) plus noise caused by Brownian motion multiplied by \( b(t, X_t) \) for some stochastic process \( X = (X_t, t \in [0, T]) \).

Next, we will state Itô’s formula which is a fundamental tool used in Itô processes. We first state it for Brownian motion.

Let \( f(t, x) \) be a function which has partial derivatives \( f_x(t, x) \), \( f_t(t, x) \) and \( f_{xx}(t, x) \) are continuous and defined \( \forall T \geq 0 \). Let \( W(t) \) be a Brownian Motion process. Then the Itô formula for Brownian motion is given by:
\[
\begin{align*}
f(T, W(T)) &= f(0, W(0)) + \int_0^T f_t(t, W(t)) dt + \int_0^T f_x(t, W(t)) dW(t) \\
&+ \frac{1}{2} \int_0^T f_{xx}(t, W(t)) dt
\end{align*}
\] (18)

\(^5\)Often times \( a(t, x) \) will be replaced by drift coefficient \( \mu_t \) and \( b(x, t) \) with \( \sigma_t, \sigma_t^2 \) being variance coefficient
which can also be stated in differential form as:

\[
df(t, W(T)) = f_t(t, W(t))dt + f_x(t, W(t))dW(t) + \frac{1}{2} f_{xx}(t, W(t))dW(t)dW(t) + f_{tW}(t, W(t))dtW(t) + \frac{1}{2} f_{xx}(t, W(t))dt\cdot dt
\]

but: \(dW(t)dW(t) = dt, dt\cdot dW(t) = 0, dW(t)dt = 0\) and \(dtdt = 0\) so:

\[
\Rightarrow df(t, W(T)) = f_t(t, W(t))dt + f_x(t, W(t))dW(t) + \frac{1}{2} f_{xx}(t, W(t))dt\cdot dt
\]

Now, we have derived an Itô formula for Brownian Motion. However, we will require a more generalized version of Itô’s formula by which we express it in terms of Itô process \(X_t\).

Let \(f(t, x)\) be a function which has partial derivatives \(f_x(t, x), f_t(t, x)\) and \(f_{xx}(t, x)\) are continuous and defined \(\forall T \geq 0\). Let \(X(t)\) be an Itô process defined in (14). Then the Itô formula for an Itô process is given by:

\[
f(T, X(T)) = f(0, X(0)) + \int_0^T f_t(t, X(t))dt + \int_0^T f_x(t, X(t))\sigma(t)dW(t)
+ \int_0^T f_{xx}(t, X(t))\mu(t)dt + \frac{1}{2} \int_0^T f_{xx}(t, X(t))\sigma^2(t)dt
\]

which can also be stated in differential form as:

\[
df(t, X(t)) = f_t(t, X(t))dt + f_x(t, X(t))dX(t) + \frac{1}{2} f_{xx}(t, X(t))dX(t)dX(t)
\]

In this text will mainly be concerned with the differential form given in (21). This general form will be used in the HJB derivation which we will explore later.

Further, if we know the stochastic process followed by \(X\) we can define some function \(G(x, t)\) to represent model (15). As there is no clear solution and often times (15) is not well defined it proves useful to differentiate \(G(x, t)\) in order to yield a solution. This leads to the main concept of this section, Itô’s lemma.

Consider a stock price process \(dS = \mu Sdt + \sigma Sdz\) for a function \(G(S, t)\) where \(dz\) is a Wiener process and \(\mu\) and \(\sigma\) can be functions of \(x\) and \(t\). Then differentiating \(G\) yields:

\[
dG = \left[\frac{\partial G}{\partial S}\mu S + \frac{\partial G}{\partial t} + \frac{1}{2} \frac{\partial^2 Q}{\partial S^2} \sigma^2 S^2\right]dt + \frac{\partial G}{\partial S}\sigma Sdz
\]

where \(dz\) is an Itô process. Hence, an Itô process is a closed process for differentiation up to the second degree. This process is commonly used in the derivation of option prices in the Black-Scholes equation.

25
4.2.1 Example: SDE representation of GBM process

We will end this section of the text with an example of computing the stochastic differential equation representation of the Geometric Brownian Motion stock model process.

Consider a Brownian Motion process $W(t)$ with associated filtration $\mathcal{F}(t)$ and let $\mu(t)$ and $\sigma(t)$ be adapted processes where $t \geq 0$. Then, for the Itô process given by:

$$X(t) = \int_0^t \left( \mu(s) - \frac{1}{2} \sigma^2(s) \right) ds + \int_0^t \sigma(s) dW(s)$$

then we can state this in differential form as:

$$dX(t) = \left( \mu(t) - \frac{1}{2} \sigma^2(t) \right) dt + \sigma(t) dW(t)$$

$$\Rightarrow dX(t) dX(t) = \sigma^2(t) dW(t) dW(t)$$

and since $dt = dW(t) dW(t) = \sigma^2(t) dt$

So if we model the price of a stock using the formula:

$$S(t) = S(0) e^{X(t)}$$

$$= S(0) e^{\int_0^t (\mu(s) - \frac{1}{2} \sigma^2(s)) ds + \int_0^t \sigma(s) dW(s)}$$

and since the stock price is non-negative:

$$\Rightarrow$$ setting $S(t) = f(X(t))$

so that: $f(x) = S(0) e^x$, $f'(x) = S(0) e^x$ and $f''(x) = S(0) e^x$

we can apply the Itô formula for an Itô process:

$$dS(t) = df(X(t))$$

$$= f'(X(t)) dX(t) + \frac{1}{2} f''(X(t)) dX(t) dX(t)$$

$$= S(0) e^{X(t)} dX(t) + \frac{1}{2} S(0) e^{X(t)} dX(t) dX(t)$$

$$= S(t) dX(t) + \frac{1}{2} S(t) dX(t) dX(t)$$

$$= \mu S(t) dt + \sigma(t) S(t) dW(t)$$

$$= S(t)(\mu(t) dt + \sigma(t) dW(t))$$
the SDE representation of a GBM process.

5 Stochastic Control

We will consider control variables of the system to be $\pi_t$, a multidimensional vector which the investor can control. $\pi_t$ represents $n$, the number of risky assets held, and $c_t$, the consumption at time $t$ which can be adjusted at any point for optimal performance given available information. That is, $\pi_t = (n_t, c_t)$ for control.

We consider $dX_t$ to follow diffusion control process:

$$dX_t = \mu(X_t, \pi)dt + \sigma(X_t, \pi)dW_t$$  \hfill (23)

through which an investor can allocate and consume their resources.

Definition 5.1 (Solution to the diffusion). For any time $s < t \in [0, T]$ we consider the solution to the diffusion process starting from $x \in \mathbb{R}$ at time $t$ to be represented by:

$$\{X_{s,x}^{s,x}, s < t < T\} \hfill (24)$$

Hence, we can define the objective function that an investor wishes to maximize over a finite period as:

$$J(x, \pi, t) := \mathbb{E} \left[ \int_t^T f(s, X_s^t, \pi_s)ds + g(X_T^t) \right] \quad \text{when } T < \infty \hfill (25)$$

and over an infinite horizon:

$$J(x, \pi) := \mathbb{E} \left[ \int_0^\infty e^{-\rho t} f(X_s^t, \pi_s)ds \right] \quad \text{when } \rho > 0$$

Where $f$ is a generalization to an investors utility function, $g$ is the generalized terminal value function and $\rho$ is the discount rate. Based on the definition of the control variables, one can express this optimal control process without effecting the results in either $\theta, \pi$ or $n$ since, $\theta_t = n_t S_t = \pi_t w_t \forall t$ where $S_t$ and $w_t$ is the stock price and wealth respectively at time $t$. 

27
5.1 Derivation of the HJB Equation

In the formal derivation of the *Hamilton-Jacobi-Bellman equation*, we will apply limit techniques and compute the derivative from first principals. Consider equation (25). The function \( J \) is what is known as the *gain function* which leads to a key definition of the *value function*, \( v(t, x) \), which is simply the supremum of the gain function over admissible control variables \( \pi = (n, c) \) in the form:

\[
J(t, x, \pi) = \mathbb{E} \left[ \int_t^T f(X^{t,x}_s, \pi, t) dt + g(X^{t,x}_T, t) \right]
\]  

\( \Rightarrow v(t, x) = \sup_{\pi} J(t, x, \pi) \) (27)

and for optimal control variables we achieve: \( v(t, x) = J(x, \pi^*) \), optimal levels \( \pi^* = (n^*, c^*) \).

It is obvious from the definition of \( v \) for fixed \( \pi = \pi_t \) that:

\[
v(t, x) \geq \mathbb{E} \left[ \int_t^{t+h} f(s, X^{t,x}_s, \pi) ds + v(t + h, X^{t,x}_{t+h}) \right] \]  

(28)

as \( v \) is defined to be the supremum over all control variables. Now, applying Itô’s formula over a small period between \( t \) and \( t + h \) when we consider \( dX_t \) to follow diffusion control process (23) we find the value function to follow the form:

\[
v(t + h, X^{t,x}_{t+h}) = v(t, x)
+ \int_t^{t+h} \left( \frac{\partial v}{\partial t} + (\mu(x, \pi)v_x + \frac{1}{2}(\sigma^2(x, \pi)v_{xx}) \right)(s, X^{t,x}_s)ds
+ \text{martingale}
\]

in accordance with §4.1.1 since the stochastic integral is a martingale. We can substitute this value for \( v(t + h, X^{t,x}_{t+h}) \) into (28) to find that:

\[
v(t, x) \geq v(t, x)
+ \mathbb{E} \left[ \int_t^{t+h} f(s, X^{t,x}_s, \pi) ds + \left( \frac{\partial v}{\partial t} + (\mu(x, \pi)v_x + \frac{1}{2}(\sigma^2(x, \pi)v_{xx}) \right)(s, X^{t,x}_s)ds
+ \text{martingale} \right]
\]
since the expected value of a martingale is zero, we simplify this result:

\[ 0 \geq \mathbb{E} \left[ \int_t^{t+h} f(s, X_{s,x}^t, \pi) ds + \left( \frac{\partial v}{\partial t} + (\mu(x, \pi)v_x + \frac{1}{2}(\sigma^2(x, \pi)v_{xx}) \right)(s, X_{s}^t) \right] \]

So, now we can divide both sides by \( h \) and send \( h \) to zero to find the derivative by first principles:

\[ \lim_{h \to 0} \frac{0}{h} = \lim_{h \to 0} \mathbb{E} \left[ \int_t^{t+h} f(s, X_{s,x}^t, \pi) ds + \left( \frac{\partial v}{\partial t} + (\mu(x, \pi)v_x + \frac{1}{2}(\sigma^2(x, \pi)v_{xx}) \right)(s, X_{s}^t) \right] / h \]

So by the assumed continuity of \( f \) and \( v \) we apply the mean-value theorem to find some existent point \( t \) giving:

\[ 0 \geq f(t, x, \pi) + \frac{\partial v}{\partial t}(t, x) + \mu(x, \pi)v_x(t, x) + \frac{1}{2}(\sigma^2(x, \pi)v_{xx}(t, x)) \]

In which we can express in any admissible combination of control variables \( \pi = (n, c) \):

\[ \Rightarrow - \frac{\partial v}{\partial t}(t, x) - \sup_{\pi} [f(t, x, \pi) + \mu(x, \pi)v_x(t, x) + \frac{1}{2}(\sigma^2(x, \pi)v_{xx}(t, x))] \geq 0 \] (29)

Now instead, if we take optimal values for \( \pi = (n, c) \) we know that the inequality expressed in (28) becomes an equality by the definition of \( v \) so now let:

- \( \pi_t^* \) be the optimal coordinate \((n^*, c^*)\) at time \( t \)
- \( c_t^* \) be the optimal level of consumption at time \( t \)
- \( n_t^* \) be the optimal number of risky assets held at time \( t \)
- \( X^* \) be the state solution starting with wealth, \( x = x_0 \) at time \( t \) with optimal control values \( \pi^* \) and \( c^* \)

\[ \Rightarrow v(t, x) \geq \mathbb{E} \left[ \int_t^{t+h} f(s, X_{s,x}^*, \pi^*) ds + v(t + h, X_{t+h}^*) \right] \]

And by (29),

\[ - \frac{\partial v}{\partial t}(t, x) - [f(t, x, \pi^*) + \mu(x, \pi^*)v_x(t, x) + \frac{1}{2}(\sigma^2(x, \pi^*)v_{xx}(t, x))] = 0 \] (30)

That is, we have found an optimal expression of an investors wealth process over time with optimal control variables. Hence, if one can derive these optimal levels then the investor will
maximize their well-being over time $T$ by optimally consuming and investing to maximize returns.

This function (30) can be expressed as:

$$-\frac{\partial v}{\partial t}(t, x) - H(t, x, v_x(t, x), v_{xx}(t, x)) = 0$$  \hfill (31)

Equation (31) is called the Hamilton-Jacobi-Bellman (HJB) equation. We find optimal control by optimizing the function $H$.

**Definition 5.2** (Hamiltonian finite-horizon case). Function $H$ in called the Hamiltonian of the associated control problem which over the finite time horizon is expressed as:

$$H(t, x, v_x(t, x), v_{xx}(t, x)) = \sup_{\pi} \left[ f(t, x, \pi) + \mu(x, \pi)v_x(t, x) + \frac{1}{2}(\sigma^2(x, \pi)v_{xx}(t, x)) \right]$$  \hfill (32)

The arguments used in this section can also be applied for the infinite-horizon case. In equation (26) we defined the objective function to be:

$$J(x, \pi) := \mathbb{E} \left[ \int_0^\infty e^{-\rho t} f(X^x_s, \pi_s) ds \right]$$

then the value function would not be dependent on time as the finite case was, so it follows that:

$$v(x) = \sup_{\pi} J(x, \pi)$$

$$\Rightarrow 0 = \rho v(x) - \sup_{\pi} \left[ f(x, \pi^*) + \mu(x, \pi^*)v_x(x) + \frac{1}{2}(\sigma^2(\pi^*)v_{xx}(x)) \right]$$

similarly, we can now define the Hamiltonian for the infinite-horizon case.

**Definition 5.3** (Hamiltonian infinite-horizon case). Function $H$ in called the Hamiltonian of the associated control problem for the infinite-horizon case is expressed as:

$$H(x, v_x(t, x), v_{xx}(x)) = \sup_{\pi} \left[ f(x, \pi) + \mu(x, \pi)v_x(x) + \frac{1}{2}(\sigma^2(\pi)v_{xx}(x)) \right]$$  \hfill (33)

### 5.2 The value function

We now have the tools to begin solving the HJB equation. We have stated that there exists a value function however at the moment we have no understanding on what it looks like. In
this section we begin to unravel this function and develop its form. Then, we will aim to
find the optimal consumption and the amount invested in risky assets over the infinite time
horizon. In doing so we can also solve the life-time consumption model for an investor.
In equation (26) we defined the objective function to be:

$$J(x, \pi) := \mathbb{E} \left[ \int_0^\infty e^{-\rho t} f(X_s^x, \pi_s) ds \right]$$

whose value function \(^6\) would not be dependent on time:

$$V(x) = \sup_{\pi} J(x, \pi)$$

Where we now substitute this general function \(f\) for the investors utility function described
in (11) giving:

$$V(x) = \sup_{n,c} I \mathbb{E} \left[ \int_0^\infty e^{-\rho t} \frac{c_1^{1-R}}{1-R} dt \right]$$

So now we consider the value function and some constant \(\gamma > 0\):

$$V(\gamma x) = \sup_{n,c} \mathbb{E} \left[ \int_0^\infty e^{-\rho t} \frac{\gamma c_1^{1-R}}{1-R} dt \right]$$

$$= \sup_{n,c} \gamma^{1-R} \mathbb{E} \left[ \int_0^\infty e^{-\rho t} \frac{c_1^{1-R}}{1-R} dt \right]$$

$$= \gamma^{1-R} V(x)$$

Taking \(w = 1\) in the preceding result means that we can immediately write down the form
of the value function given by:

$$V(x) = \gamma^{-R} \frac{x^{1-R}}{1-R}$$

for some constant \(\gamma > 0\). Further, discounting the value function back to time \(t\) we find that:

$$V(t, x) = \sup_{n,c} \mathbb{E} \left[ \int_t^\infty e^{-\rho s} \frac{\gamma c_1^{1-R}}{1-R} ds \mid w_t = w \right]$$

$$= e^{-\rho t} V(w)$$

31
so now,

\[ V(t, x) = e^{-\rho t} \sup_{n, c} \mathbb{E} \left[ \int_0^\infty e^{-\rho t} \left( \gamma c_t \right)^{1-R} \frac{1}{1 - R} dt \right] \]

\[ = e^{-\rho t} \gamma^{-R} u(x) \]  

(34)

Now that we know the suspected form of the value function we can optimize the HJB (31) over \( \theta \) and \( c \) separately to find a solution to the Merton portfolio problem.

**Definition 5.4** (Merton Portfolio Problem). The *Merton Portfolio problem* is a question of how much to consume and how much wealth to allocate to \( d \) risky assets of proportion \( \sum_{i=1}^{d} \pi_M^i \) and the remaining to the risk-free asset \( 1 - \sum_{i=1}^{d} \pi_M^i \). This problem is addressed in a continuous time environment with subscript \( M \) representing an optimal solution.

### 5.3 Optimization over \( \theta \)

We start the optimization process with respect of \( \theta \). The process below follows provided \( V(t, w) \) has a strictly positive wealth process and that \( V(t, w) \) is strictly concave in \( w \). Given these assumptions (which can be proven) we can now use the general Hamiltonian for the infinite-horizon case as described in (33) to find an investors optimal allocation of resources.

**Remark.** In this section we will substitute initial wealth of the process \( X_0 = x \) for \( X_0 = w \) and will optimize \( \theta_t^i = n^i_t S_t^i \) which will not harm any of the statements outlined in the stochastic control section which were mentioned in terms of \( n^i_t \).

Consider the general form of the Hamiltonian equation given as:

\[ 0 = \sup_{\pi} [f(x, \pi) + \mu(x, \pi)v_x(x) + \frac{1}{2} \sigma^2(\pi)v_{xx}(x)] \]

At this point it will be useful to begin to lose some of the generality to arrive at a meaningful equation. To do so we will replace \( x \) with wealth \( w \) as described in the remark above. Similarly to the previous section, we will also substitute general function \( f \) for the CRRA
utility function as described in (10). Lastly, we will consider one component of the optimal value \( \theta \) stated above. So now we can restate generalized equation (33) in the form of:

\[
0 = \sup_{\theta} \left[ u(t, c) + V_t + V_w (rw + \theta (\mu - r) - c) + \frac{1}{2} |\sigma^T \theta|^2 V_{ww} \right]
\]

and after some simple algebra:

\[
\Rightarrow (\sigma \sigma^T) \theta V_{ww} = - (\mu - r) V_w
\]

so we can solve for optimal \( \theta, \theta^* \):

\[
\theta^* = - (\sigma \sigma^T)^{-1} (\mu - r) \left( \frac{V_w}{V_{ww}} \right)
\]

and by (34) we believe that the value function should follow the form \( e^{-\rho t} \gamma^{-R} u(w) \)

\[
\Rightarrow \theta^* = - (\sigma \sigma^T)^{-1} (\mu - r) \left( \left[ \frac{\partial V}{\partial w} \right] \right)^{-1}
\]

\[
= - (\sigma \sigma^T)^{-1} (\mu - r) (1 - R) \left( \frac{w^{-1} - 1}{1 - R} \right) (-R)^{-1} \left( \frac{1}{w^{-R} - 1} \right)
\]

\[
= - (\sigma \sigma^T)^{-1} (\mu - r) (-R)^{-1} \left( \frac{w^{-R}}{w^{-R} - 1} \right)
\]

\[
= - (\sigma \sigma^T)^{-1} (\mu - r) (-R)^{-1} (w)
\]

This very simple form has lead us to an optimal value for \( \theta^* \). We can restate this equation with another useful representation. Recalling (9) that at time \( t \) asset \( i \) represents proportion of wealth \( \pi_i^t \) will have a cash value given by \( \theta_i^t = \pi_i^t w_t \) where \( \sum_{i=1}^d \pi_i^t = 1 \). Hence, by (36) we can find an optimal allocation of wealth to asset \( i \) which is proportional to the investors wealth at any time \( t \). That is,

\[
(\theta_i^t)^i = w_t \pi_i^t
\]

We can substitute this into the above giving:

\[
\pi_M = - (\sigma \sigma^T)^{-1} (\mu - r) (-R)^{-1}
\]

Equation (38) suggests an optimal solution to part of the Merton Portfolio problem. The Merton Portfolio problem aims to understand how much wealth to allocate to \( d \) risky assets of proportion \( \sum_{i=1}^d \pi_i^M \) and the remaining to the risk-free asset \( 1 - \sum_{i=1}^d \pi_i^M \). Through optimal control, this is exactly what we have provided. The remaining is consumption.
5.4 Optimization over $c$

Next, we look to optimize the Hamiltonian (33) with respect to consumption $c$. This method is less straightforward than the previous optimization problem in which we will need to introduce another function - a convex dual function of $u$ given by:

$$\tilde{u}(y) = \sup_c \{u(c) - cy\}$$

where $u(x)$ is the CRRA utility function. Setting $\tilde{R} = R^{-1}$:

$$\Rightarrow \tilde{u}(y) = -\frac{y^{1-R}}{1-R}$$

Hence, we can optimize over $c$ and involve $t$ for discounting to compute:

$$\sup_c \{u(t,c) - cV_w\} = \sup_c \{e^{-\rho t}u(c) - ce^{-\rho t}\beta V_w\}$$

$$= e^{-\rho t} \sup_c \{u(c) - ce^{\rho t}V_w\}$$

$$= e^{-\rho t} \tilde{u}(e^{\rho t}V_w)$$

and once again, by (34) we believe that the value function should follow the form $e^{-\rho t\gamma - R}u(w)$

$$\sup_c \{u(t,c) - cV_w\} = e^{-\rho t} \tilde{u}(e^{\rho t}V_w)$$

$$= e^{-\rho t} \tilde{u}((w\gamma)^{-R})$$

and by the definition of $\tilde{R}$ we evaluate the function $\tilde{u}((w\gamma)^{-R})$

$$= -e^{-\rho t} \frac{(w\gamma)^{1-R}}{1-R}$$

and after some algebra

$$\Rightarrow \sup_c \{u(t,c) - cV_w\} = e^{-\rho t} \frac{R}{1-R} (w\gamma)^{1-R}$$

Hence, we have derived an optimal level for $c$ which is relative to parameter $(w\gamma)$. That is, we should select an optimal $c$ which is proportional to wealth, $w$:

$$c^* = w\gamma$$
which can be expressed similar to (38) as the optimal consumption solution to the Merton problem\(^7\) with notation:

\[ c^* = w\gamma_M \]  \hspace{1cm} (42)

5.5 Optimization over \(\theta\) and \(c\)

At this point we have derived optimal levels of consumption and proportions of wealth for each asset holding, \((\theta^*, c^*)\). So, we look to optimize the HJB equation given these constraints to solve the Merton problem. Hence we compute \(V_t, V_w\) and \(V_{ww}\) and by (34) the form of the value function we substitute the optimal values into (31) the HJB to find:

\[
0 = \left[ e^{-\rho t} \frac{R}{1 - R} (w\gamma)^{1-R} - \rho \gamma^{-R} u(w) + e^{-\rho t} R \gamma^{-R} w^{-R} + \frac{1}{2} e^{-\rho t} \gamma^{-R} w^{-R} |\sigma^{-1}(\mu - r)|^2 / R \right]
\]

substituting \(u(w)\) and simplifying:

\[
\Rightarrow 0 = \frac{e^{-\rho t} w^{1-R} \gamma^{-R}}{1 - R} \left[ R\gamma - \rho - (R - 1) \left( r + \frac{|\sigma^{-1}(\mu - r)|^2}{2R} \right) \right]
\]

\[
\Rightarrow \gamma = R^{-1} \left[ \rho + (R - 1) \left( r + \frac{|\sigma^{-1}(\mu - r)|^2}{2R} \right) \right] \hspace{1cm} (43)
\]

Now, (43) has identified the constant \(\gamma\) for which the value functions’ form was stated in equation (34). Therefore, the value function in the Merton problem must be:

\[
V(t, w) = V_M(w) = \gamma^{-R} u(w) \hspace{1cm} (44)
\]

\(^7\)A subscript \(M\) refers to an optimal constraint to the Merton portfolio problem.
5.6 Formal statement of the problem

In summary of §5 we restate the problem and how we achieve optimal control.

The portfolio problem

An investor aims to find a value function $V(w)$ that will maximize the expected value of their utility function for optimal values of $(n, c)$ which satisfy initial wealth $w_0$ are given by $(n^*, c^*)$. That is, an investor wishes to optimize:

$$V(w) = \sup_{n,c} \mathbb{E} \left[ \int_0^\infty e^{-\rho t} \frac{e^{c_t^{1-R}}}{1-R} dt \right]$$  \hspace{1cm} (45)

Where,

$$c^* = w\gamma_M$$  \hspace{1cm} (46)

$$(\theta^*_t)^i = w_t\pi^i_M$$  \hspace{1cm} (47)

$$\pi_M = -(\sigma \sigma^T)^{-1}(\mu - r)(-R)^{-1}$$  \hspace{1cm} (48)

$$\gamma_M = R^{-1}\left[ \rho + (R - 1) \left( r + \frac{|\sigma^{-1}(\mu - r)|^2}{2R} \right) \right]$$  \hspace{1cm} (49)

So that the investor will invest in asset $i$ proportional to their wealth $(\theta^*_t)^i$, consume proportional to their wealth $c^*$ all in proportion to the constraints given by $\pi_M$ and $\gamma_M$ respectively.

6 Simulations: 2

Now that we have found a solution to the Merton problem we can simulate the performance of this optimal allocation of resources. For simplicity we will ignore consumption in our model and only consider the proportion of wealth invested in risky assets with the remainder being invested at the risk-free rate.

$$T = 50 \quad X_0 = \$100,000, \quad \sigma = 0.35, \quad r = 0.05, \quad S_0 = 100, \quad R = 2, \quad \mu = 0.15$$

First, we will examine the performance of optimal investment versus investing solely at the risk-free rate. In this simulation we assume for there to be $d = 10$ stocks each with an initial
stock price of 100. That is, $S_0^i = 100$, $\forall i = 1, \ldots, d$. We invest a proportion of wealth of $\pi_M$ equally into each stock which will be adjusted to the optimal level based on the solution to the Merton problem for each change in time. It is important to note that this division of wealth between the two types of assets is based on an individual's risk-aversion parameter $R$.

![Simulated performance of portfolio with optimal wealth allocation](image)

Figure 5: Simulated performance of portfolio with optimal wealth allocation

The previous simulation demonstrates how a portfolio of uncorrelated stocks will generate a positive return which approaches its expected value. Next, we look to see the effects of investing optimally in a combination of the risk-free asset and a single risky asset. We will also simulate the effects of two investors with different risk aversion parameters and their respective returns. In this simulation we use the same coefficients as listed above with the exception of the second investor being more risk tolerant.
Next, we can look back on the concept of geometric Brownian Motion. In this model we developed $d$ independent stock processes. However, this is unrealistic in the real world as market sectors tend to move together.

Correlated Geometric Brownian Motion processes can be developed by correlating the Brownian Motion component $W_t$ between stocks. The process remains the same except we compute the Cholesky factorization on the correlation matrix and multiply this by the uncorrelated Brownian Motion Process $W_t$. Below is a simulation of the correlated stock processes for different correlation coefficients.
7 Simulations: 3 - A brief study on five Technology Giants

The five best performing companies on the stock market are coined to be the FAANG stocks. This acronym represents companies: Facebook, Apple, Amazon, Netflix and Google (Alphabet). In this simulation we will analyse historical returns to develop a simulation using a correlated GBM process. Below is some data using historical returns over the past 1-year period.
Now we can simulate the asset prices over the next one-year period using these historical figures as a proxy.
Lastly, we would like to simulate an optimal portfolio for an investor who wishes to invest solely in FAANG stocks and the risk-free asset. We will use the one year US treasury bill rate to proxy the risk-free rate and one year historical data for each stock mentioned above. We will apply the optimal solution to the Merton portfolio problem for an investor with risk aversion parameter $A = 2$. Below is the results for an optimal composition of stocks/T-bills:

![Figure 10: Optimal wealth allocation (≈ 45% risky, ≈ 55% T-bill)](image)

We can now simulate these stocks and view the results on the investors portfolio.
Figure 11: Simulating optimal wealth composition portfolio process over one year

Which produced the annual returns below:

Figure 12: Simulated optimal wealth composition portfolio return (%)

Hence, we have found that even given the volatile technology market, the investor managed to outperform the risk-free return rate with a risk aversion parameter of 2 by around 4.88%.
References


Appendices

A MATLAB Code

% Brownian Motion
clear all
N = 1000;
randvalsW = randn(1,N);

% Generate W
W = cumsum(sqrt(1/N) * randvalsW); % Sum over increments

% paths
plot([0:1/N:1],[0,W])
title('Simulated path of Brownian Motion over [0,1]

% Geometric Brownian Motion
clear all
format long
% user choice of value of p=percentage return
p=115763/100000;

mu=[0.5 0 -0.5];
T=3;
N = T*365; % number of paths
n = T*365; % number of MC simulations
Snot=100; % initial stock value
sigma=0.25;
dt=T/N;

for idx = mu
    S = Snot + zeros( 1, N );
    for k = 1:N
        for i=1:N
            S(k+1,i)=S(k, i)*exp((idx−0.5*sigma^2)*dt ... + sigma*sqrt(dt)*randn);
        end
    end
end

%computes mean path of n simualtions
meanrow_S = round(mean(S'), 0);

txt = ['\mu = ', num2str(idx)];
plot([0:dt:T], meanrow_S, 'LineWidth' ... ,2,'DisplayName',txt)
hold on
end
title(['Average path of N = ', ... num2str(1000), ' simulated stock prices'])
legend('Location', 'NorthWest');
pointText = ['\sigma = ', num2str(sigma)];
text(0.25, 350, pointText, 'VerticalAlignment', 'top');
hold off
legend show

%Correlated Geometric Brownian Motion
clear all;

T=1;
N = T*365; %number of paths
n = T*365; %number of MC simulations
Snot=100; %initial stock value
Xnot=100000; %initial portfolio value
riskfree = [Xnot];
portstocks = 2;
X = Xnot + zeros( 1 , portstocks );
Xrisk = Xnot + zeros( 1 , portstocks );
S = Snot + zeros( 1 , portstocks );
r=0.05;
rho = 0.2;
rComp=((1+r)^(1/365));
m=157;
mu=0.15;
sigma=0.35;
dt=T/N;

corr = [1 -0.99; -0.99 1];
choles = chol(corr);
uncor = randn(N, size(corr,2));
correl = uncor*choles;

%generate stock prices
for k = 1:N
    for i=1:portstocks
        S(k+1,i)=S(k,i)*exp((mu-0.5*sigma^2)*dt ... 
        + correl(k,i)*sigma*sqrt(dt));
end
end

subplot(3,1,1);

plot([0:dt:T],S(1:N+1,1));
hold on;
plot([0:dt:T],S(1:N+1,2));
title('Correlation = -0.99')
corr = [1 0.5; 0.5 1];
choles = chol(corr);
uncor = randn(N,size(corr,2));
correl = uncor*choles;

%generate stock prices
for k = 1:N
    for i=1:portstocks
        S(k+1,i) = S(k,i)*exp((mu-0.5*sigma^2)*dt ... + correl(k,i)*sigma*sqrt(dt));
    end
end
subplot(3,1,2);

plot([0:dt:T],S(1:N+1,1));
hold on;
plot([0:dt:T],S(1:N+1,2));
title('Correlation = 0.5')
corr = [1 0.99; 0.99 1];
choles = chol(corr);
uncor = randn(N,size(corr,2));
correl = uncor\*choles;

%generate stock prices
for k = 1:N
    for i=1:portstocks
        S(k+1,i)=S(k,i)*exp((mu-0.5*sigma^2)*dt . . .
        + correl(k,i)*sigma*sqrt(dt));
    end
end

subplot(3,1,3);
plot([0:dt:T],S(1:N+1,1));
hold on;
plot([0:dt:T],S(1:N+1,2));
title('Correlation = 0.99')

%FAANG stocks
clear all;
coreldata = readtable('FAANGdata.xlsx','Range','B1:F6');
vardata = readtable('FAANGdataSigma.xlsx','Range','A1:E2');
mudata = readtable('FAANGdataDrift.xlsx','Range','A1:E2');
coreldata = table2array(coreldata);
sigma = table2array(vardata);
mu = table2array(mudata);
T=1;
N = T*365; %number of paths
n = T*365; %number of MC simulations
Xnot=100000; %initial portfolio value
riskfrees = [Xnot];
portstocks = 5;
X = Xnot + zeros(1, portstocks);
Xrisk = Xnot + zeros(1, portstocks);
snotFB = 191.49;
snotAAPL = 204.30;
snotAMZN = 1950.630;
snotNFLX = 374.85;
snotGOOG = 1272.18;
S = [snotFB snotAAPL snotAMZN snotNFLX snotGOOG];

r = 0.02421;
rho = 0.2;
rComp = ((1+r)^(1/365));
m = 157;
dt = T/N;

corr = coreldata;
choles = chol(corr);
uncor = randn(N, size(corr,2));
correl = uncor*choles;

% generate stock prices
for k = 1:N
    for i = 1:portstocks
        S(k+1,i) = S(k,i)*exp((mu(i) - 0.5*sigma(i)^2)*dt ... + correl(k,i)*sigma(i)*sqrt(dt));
    end
end
```matlab
% subplot (3,1,1);
subplot (5,1,1);
plot ([0:dt:T],S(1:N+1,1), 'b');
title ('FB')
hold on;
subplot (5,1,2);
plot ([0:dt:T],S(1:N+1,2), 'k');
title ('AAPL')
subplot (5,1,3);
plot ([0:dt:T],S(1:N+1,3), 'color', [0.9100 0.4100 0.1700]);
title ('AMZN')
subplot (5,1,4);
plot ([0:dt:T],S(1:N+1,4), 'r');
title ('NFLX')
subplot (5,1,5);
plot ([0:dt:T],S(1:N+1,5), 'g');
title ('GOOG')

% % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % % %
%optimal FAANG portfolio
clear all;
coredata = readtable ('FAANGdata.xlsx', 'Range', 'B1:F6');
vardata = readtable ('FAANGdataSigma.xlsx', 'Range', 'A1:E2');
mudata = readtable ('FAANGdataDrift.xlsx', 'Range', 'A1:E2');
coredata = table2array (coredata);
sigma = table2array (vardata);
mu = table2array (mudata);
T=1;
N = T*365; %number of paths
n = T*365; %number of MC simulations
```
Xnot=100000; \% initial portfolio value
riskfree = [Xnot];
portstocks = 5;
X = Xnot + zeros( 1, portstocks );
Xrisk = Xnot + zeros( 1, portstocks );
numShares = [];
snotFB = 191.49;
snotAAPL = 204.30;
snotAMZN = 1950.630;
snotNFLX = 374.85;
snotGOOG = 1272.18;
S = [snotFB snotAAPL snotAMZN snotNFLX snotGOOG];

r = 0.02421;
rho = 0.2;
rComp = ((1 + r)^(1/365));
m = 157;
dt = T/N;

A = [];
Port = [];
Tradingtimes = [];
TradingIndicator = [];
pi = [];
R = 2;
for i = 1:portstocks
    pi(i) = (sigma*sigma')^(-1)*(mu(i) - r)*(R)^(-1);
end;
for i = 1:portstocks
    numShares(i) = ((X(i))*(pi(i))/S(i)); \% initial number of shares
    %initial number of shares

end;
invested=(Xnot)*(1−sum(pi))+zeros(1,1);

genvals=[];

corr = coreldata;
choles = chol(corr);
uncor = randn(N,size(corr,2));
correl = uncor*choles;

%generate stock prices
for k = 1:N
    for i=1:portstocks
        S(k+1,i)=S(k,i)*exp((mu(i)−0.5*sigma(i)^2)*dt ... + correl(k,i)*sigma(i)*sqrt(dt));
    end
end

subplot(7,1,1);
plot([0:dt:T],S(1:N+1,1),'b');
title('FB')
hold on;
subplot(7,1,2);
plot([0:dt:T],S(1:N+1,2),'k');
title('AAPL')
subplot(7,1,3);
plot([0:dt:T],S(1:N+1,3),'color',[0.9100 0.4100 0.1700]);
title('AMZN')
subplot(7,1,4);
plot ([0:dt:T],S(1:N+1,4),'r');
title ('NFLX')
subplot(7,1,5);
plot ([0:dt:T],S(1:N+1,5),'g');
title ('GOOG')

%Generate adjustment periods
for i = 1:m-1
    Tradingtimes=[Tradingtimes,round (i*T/m, 15)];
end
for i = 1:N+1
    if mod(i,round(N/m)) == 0
        TradingIndicator(i)= 1;
    else
        TradingIndicator(i)= 1;
    end
end

for k = 1:N+1
    for i=1:portstocks
        if TradingIndicator(k)==1 && k>1
            numShares(k,i)= (sum(X(k-1,:),'all')) ...*(pi(i))/S(k-1,i);
            invested(k,1)=round(sum(X(k-1,:),'all') ...*(1-sum(pi))*rComp,2);
        elseif k==1
            numShares(k,i)= numShares(k,i);
            invested(k,1) = round(invested(k,1),0);
        else

```
numShares(k, i) = numShares(k−1, i);

invested(k, 1) = round(invested(k−1, 1)*rComp, 0);
end

X(k, i) = round(round(S(k, i), 2)*numShares(k, i) + invested(k, 1)/portstocks, 2);
end
end

X2 = sum(X, 2);
S2 = S*Xnot/100;

% risk free portfolio
for i = 2:N+1
    riskfree(i, 1) = riskfree(i−1, 1)*rComp;
end

plot([0:dt:T], riskfree(1:N+1, 1))

subplot(7,1,6);
plot([0:dt:T], X2(1:N+1, 1));
hold on;
plot([0:dt:T], riskfree(1:N+1, 1))
title(’Optimal and risk−free’)

Port_return = (X2(N+1, 1)−X2(1, 1))/X2(1, 1)*100;

riskFr_return = (riskfree(N+1, 1)− ... riskfree(1, 1))/riskfree(1, 1)*100;
Stock_return = [];
for i = 1:portstocks
    Stock_return(i) = (S(N+1, i)−S(1, i))/S(1, i)*100;
end
for i = 1:portstocks
    plot([0:dt:T], S2(1:N+1, i))
    hold on
end
legend('Optimal Portfolio', 'Risk Free portfolio')

% c = categorical({'FB', 'AAPL', 'AMZN', ...
                   'NFLX', 'GOOG', 'Risk-Free'});
% b = bar(c, [pi (1-sum(pi))])

% c = categorical({'FB', 'AAPL', 'AMZN', 'NFLX', ...
                   'GOOG', 'Risk-Free', 'Portfolio-return'});
% b = bar(c, [Stock_return riskFr_return Port_return])