*Department of Economics and Finance*

*Chair of Mathematical Finance*

# Portfolio Optimization in Continuous Time

SUPERVISOR

Prof. Sara Biagini

CANDIDATE

Eugenia Marzano 216031

ACADEMIC YEAR

2018/2019

Table of contents

# 1. Introduction

With the present thesis we will deal with the portfolio optimization problem. The concepts of portfolio optimization and diversification have been instrumental in the development and understanding of financial markets and financial decision making. The major step toward quantitative management of portfolio was made by Harry Markowitz in his paper "Portfolio Selection" published in 1952 in the Journal of Finance. When Harry Markowitz, William Sharpe and Merton Miller were awarded in 1990 of the Nobel Prize in Economics, it was clear the need of a new scientific discipline, the "theory of finance. The theory of finance has become increasingly mathematical, to the point that problems in finance are now driving research in mathematics. The theory, popularly referred to as Modern Portfolio Theory, gave a response to the crucial question of how an investor should allocate funds among the possible investment choices. Markowitz proposed that agents ought to consider risk and return together and then decide the allocation of funds among investment alternatives based on the trade-off between them. Markowitz's ideas pose the base of what is now called mean-variance optimization and Modern Portfolio Theory (Hull et al.,2002)

In 1969, Robert Merton published "Lifetime Portfolio Selection Under Uncertainty: The Continuous-Time Case", here Robert Merton introduced stochastic calculus into the study of finance. He presented another approach to deal with portfolio optimization not subject to the static nature of Markowitz mean-variance approach. Specifically, we will focus on the Merton Problem.

This thesis is divided into three main sections: First section aims at giving an introduction of probability and stochastic processes and some notations required to understand the following sections. In particular, Markov Processes will be introduced as well as Geometric Brownian Motion and its property.

In the second section we deal with portfolio theory, given a brief illustration of classic Markowitz theory, that founded Modern Portfolio Theory and then enter in a specific discussion about the Merton Problem that is, optimal investment and consumption problem. In brief, let's take into account a financial market in which we have only a risk-free asset and a risky asset. In the former the price grows at a fixed rate, while in the latter the price follows a Geometric Brownian motion. An investor can allocate his wealth in two different ways: consumption and investment. Since the agent wants to maximize his expected utility from intermediate consumption and terminal wealth, he had to decide how much to consume and how to allocate his wealth between the risky asset and risk-free one. Robert Merton in his paper formulated the optimal investment and consumption problem as a stochastic optimal control problem. As a result, it leads to the Hamilton-Jacobi-Bellman equation, a fully non linear partial differential-equation (PDE). (Merton, 1992).

In the third section we analyze the complex Robust Merton Problem. Robust optimization is the mathematical discipline that takes into account the uncertainty of the problem parameters and solve the underline problem providing a solution which is the best response to data uncertainty given that uncertainty in the parameters may have a great impact on the final result. It is a technique that addresses the same type of problems as stochastic programming does but present a different approach to handling data uncertainty. It used to make relatively general assumptions on the probability distributions of the uncertain parameters so as to preserve the computational tractability. *"In robust optimization, one makes the problem well-defined by assuming that the uncertain parameters vary in a particular set defined by one's knowledge about their probability distributions, and then takes a worst-case (max-min) approach: find portfolio weights such that the portfolio return is maximized even when the vector of realizations for the asset returns takes its "worst" value over the uncertainty set."*(Fabozzi et al., 2007). In order to extend the Merton Problem to a Robust Merton Problem, we will begin explaining the portfolio optimization framework, and then we will apply it to the case of an agent that is diffident about mean and volatility returns. We will see that the result provided is based on a max-min Hamilton-Jacobi-Bellman-Isaacs, a partial differential equation that is central to optimal control theory.

# 2. Stochastic processes

## 2.1 Probability model

In order to formulate mathematics models in finance we need the stochastic calculus that is based on probability theory. For this reason, the scope of this chapter is to recall some fundamental concepts of this theory.

A variable whose value changes over time in an uncertain way is said to follow a stochastic process. Stochastic process can be divided into discrete-time processes and continuous-time processes. In the former, the value of the variable can change so it can assume different values only in determined points, In the latter, the value can change at any point.

Stochastic process can also be divided in continuous-variable and discrete variables. In the former the variable can assume all the values that are in a certain range, in the later the variable can assume only a finite number of values.

Stocks, in particular, follow a continuous time and stochastic process. For example, we know stocks are not traded continuously, they are traded when the market is open, and stocks we know can take discrete values. (Borrelli, 2012).

Nevertheless, reasoning in continuous terms is very important in achieving our result.

## 2.2 The sample space of a probability model

To elaborate probabilistic model first of all we need to fix a probability space. To give a rigorous definition, we will recall some definitions.

**Definition 2.1** We define Probability Space the measure space ($\Omega$,$\mathscr{F}$,$\mathbb{P}$) where:

$\Omega$ is the set of all possible outcomes for some experiment.

$\mathscr{F}$ is a sigma-algebra of subsets of $\Omega$, called events, which are relevant to us and which we would like to measure. It is defined as sigma-algebra and it represents the historical but not future ($\mathscr{F}$ is non-anticipative) information available on our stochastic process.

$\mathbb{P}$ is a definite measure on $\mathscr{F}$, called probability measures such that $\mathbb{P}(\Omega) = 1$.

$\mathbb{P}$ must have properties that:

- $\mathbb{P}: \mathscr{F} \rightarrow [0,1]$
- $\mathbb{P}(\Omega) = 1$ (normalization property)
- $\mathbb{P}$ is countably additive: if $(A_n)_n$ is a sequence of disjoint events, then:

$$\mathbb{P}(\sqcup_n A_n) = \sum_n \mathbb{P}(A_n)$$

These axioms imply that if $A^c$ is the complement of A, then:

$$\mathbb{P}(A^c) = 1 - \mathbb{P}(A)$$

And the principle of inclusion and exclusion:

$$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$$

Even if A and B are not disjoint.

Another way to describe a probability, given a probability space, is that of associate at each state a real number or $n$ real numbers, so to define a value function in $\mathbb{R}$ o in $\mathbb{R}_n$ on the set of the states.

If a function satisfies what we called an appropriate hypothesis it is called a random variable. Such characteristics of the probability space are strictly linked to the specific topic we are dealing with and, thus, vary with different processes meaning different ($\Omega$,$\mathscr{F}$,$\mathbb{P}$). The basic concept in probability theory is that of a random variable. Let's start by considering a real random variable $X$.

A random variable is a function of the basic outcomes in a probability space.

**Definition 2.2**. *A real random variable on* ($\Omega$,$\mathscr{F}$) *is a function on* $\Omega$ *which takes values in* $\mathbb{R}$*:*

$X : \Omega \rightarrow \mathbb{R}$

*and is* ***$\mathscr{F}$-measurable****. This means that the counter-image of any half line* $(-\infty,x]$ *is an event:*

$\{X \leq x\} \in \mathcal{F}$ *for all* $x \in \mathbb{R}$ .

The information we possess up to a certain point x generates what is called a sigma-algebra, which we define as follows:

$$\sigma(X) := \sigma(\{X \leq x\} \mid x \in \mathbb{R}$$

all the events that can be expressed in terms of X, for example $\{a \leq X \leq b\}$ belong to $\sigma(X)$.

## 2.3 Stochastic processes

We begin by emphasizing that while in deterministic processes we study a phenomenon that depends on time, of which we are able to predict the exact evolution over time, in order to describe those phenomena whose evolution is influenced by random events classical analysis is no longer adequate and it is necessary to introduce the stochastic processes, based on probability theory (Gallagher, 2013).

Particularly in the financial sector, it is not possible to accurately predict the future price of a given risky title, for example a share, knowing its past history, because this is influenced by randomness. Indeed, unpredictable events, such as the failure of a company, the collapse of a government, a terrorist action, can produce considerable fluctuations in the Stock price on the stock exchange. Due to frequent and intense variations, due to random events, the function that associates the value of an action to the variable $t$ is not derivable and therefore cannot be a solution of an ordinary differential equation. (Borrelli, 2012).

In mathematical terms a stochastic process is a sequence of real random variables, defined on the sample space, from ($\Omega$, $\mathscr{F}$) to $\mathbb{R}$.

In a filtered space $\Omega$, $(\mathcal{F}_t)_{t \leq T}, \mathbb{P})$ , if at time $t$ we know the value of a real-valued stochastic process $S = (S(t))_t$, then $S(t)$ is $F$-measurable, or, in other words, $S$ is adapted to the filtration.

For $S$ to be an adapted process the following conditions must hold:

For any fixed time $t$,

$$S(t) : \Omega \to \mathbb{R}$$

For all fixed reals $x$, the set $\{S(t) \leq x\}$ belongs to $\mathcal{F}_t$ .

## 2.4 Expected values

The expectation for a random variable $X$, is essentially the average value it expected to take on. So, the expected value E[X] , also known as the *mean*, is calculated as the weighted average of the outcomes of $X(\omega)$. This leads to the defining formula: $\mathrm{E[X]} = \sum_i x_i\, \mathbb{P}(\mathrm{X} = x_i)$ for a discrete random variable;

For continuous random variables, with density $p_X$, then $\mathbb{P}(x < X \leq x + dx) = p_X(x)dx$ and the expectation is:

$$\mathrm{E[X]} = \int x\, p_x\, (x) dx$$

Note that the distribution, and thus the expectation of a random variable, strongly depends upon $\mathbb{P}$. A change in our probabilistic views results in a change of distribution and expectation, for all the random $X$ on **Ω**. Expectation is a linear operation, meaning that the expectation of a linear combination is the linear combination of the expectations:

$$E[aX + bY] = aE[X] + bE[Y]$$

So, an expectation over a linear combination of X and Y may be computed without knowing the joint distribution $t$ of the two variables. The same cannot be said, for example, for the computation of $E[XY]$. In this case a covariance matrix is required, or, at most, assumptions over the independence of the two variables. Assume we have a continuous random variable, $X$, and function of such variable, $Y$. Then we can write:

$$Y = g(X)$$

In some case it might be useful to compute the expectation of $Y$ based on our knowledge of the distribution of $X$. There is no rule guaranteeing that $Y$ necessarily has a density. Whether or not this is the case clearly depends on $g$. If we were to face a $g$ that is of the Bernoulli type, for instance, this wouldn't be the case. But, given a $Y$ is invertible and differentiable, with $g' \neq 0,$ we can state that has density, $p_Y$ , with:

$$p_Y(y) = p_X(g^{-1}(y)) \frac{1}{|\, g'(g^{-1}(y))\,|}$$

We take now an expected value of $Y$ and by means of substitution we arrive to the following:

$$E[Y] = E[g(X)] = \int y p_X(g^{-1}(y)) \frac{1}{\mid g'(g^{-1}(y)) \mid} dy = \int g(x) p_X(x) dx$$

bearing in mind that by definition $x = g^{-1}(y)$ .

Note, moreover, that this last formula, being only dependent on $x$, is always valid and applicable to $Y$, even when it does not have a density.

## 2.5 Independence

Two random variables are independent if the joint density factors into the product of their marginal densities:

$$p_{(x,y)} = p_X(x) p_Y(y)$$

For all $x$,$y$.

If $X$ and $Y$ are independent, then:

$$E[XY] = E[X]E[Y]$$

And they are also uncorrelated:

$$E[(X - E[X])(Y - E[Y])] = 0$$

We can state that if two random variables are independent, then they are also uncorrelated. Independence always implies uncorrelation, but the opposite is not true, as two variables that are uncorrelated are not necessarily independent. (Block, 2008).

## 2.6 Conditional expectations

We now introduce a very subtle concept: conditional expectation. The simplest definition is expectation of a random variable X conditioned on an event.

Conditional expectations are more precise in making a guess over a random variable than expected values as they allow us to take into account more information that may be useful for the goal.

Some basic properties of conditional expectation are:

- $E\big[E[Y \mid X]\big] = E[Y]$
- Additivity: $E[Y_1 + Y_2] \mid X] = E[Y_1 \mid X] + E[Y_2 \mid X]$
- random variables known when X is known can be considered as constants and taken out of the expectation, $E[f(X)Y \mid X] = f(X)E[Y \mid X]$

- if two variables $X$ and $Y$ are independent we can state: $E[Y \mid X] = E[Y]$ *and* $E[X \mid Y] = E[X]$

Now we consider the case in which $\mathcal{F}_T$ ,the sigma algebra of a filtered space, represents the information over a relative specific event at time $t < T$ so if we have a set of information $\mathcal{F}_1$ our best guess is to find the conditional expectation given the set of information we already have.

So, consider:

$$E[Y|\mathcal{F}_{t1}]$$

In this case the value of $Y$ is known at time $t_2$, and this is our best guess at time $t_1$ given the information we have. And $E[Y|\mathcal{F}_{t1}]$ is a random variable itself, known, at least at some time $T$.

When information about the process $Y$ are absent, we define the trivial sigma algebra $\mathcal{F}_0$ as:

$$\mathcal{F}_0 = \{0, \Omega\}$$

Given that the information in the bracket are constant we can say:

$$E[Y \mid \mathcal{F}_0] = EW[Y \mid c] = E[Y]$$

It's now useful to list some properties of conditional expectations, taking two random variables, $Y$ and $W$, known, at least, at some time $T$. By fixing a timeline such that $0 \leq t_0 < t_1 < t_2 \leq T$ we have:

- $E\left[E[Y \mid \mathcal{F}_{t_1}]\right] = E[Y]$
- *if Y is known by time $t_1$,* $E[Y \mid \mathcal{F}_{t_1}] = Y$
- *additivity:* $E[Y + W|\mathcal{F}_{t_1}] = E[Y|\mathcal{F}_{t_1}] + E[W|\mathcal{F}_{t_1}]$
- *for any Z known at time $t_1$* ,$E[ZY|\mathcal{F}_{t_1}] = ZE[Y|\mathcal{F}_{t_1}]$
- *if Y is independent of* $\mathcal{F}_{t_1}$, *then* $E[Y|\mathcal{F}_{t_1}] = E[Y]$ *which is constant.*
- *tower law:* $E[Y|\mathcal{F}_{t_1}] = E[E[Y|\mathcal{F}_{t_1}]|\mathcal{F}_{t_0}]$; *this means that our best prediction at time $t_0$ can be made dir ectly or through an intermediate step, which is computing first the best prediction of Y for $t_1$ and then for $t_0$.*

## 2.7 Martingales

The concept of a martingale plays an important role on the modern theory of probability and in theoretical finance. By intuition, a stochastic process behaves as a martingale if his trajectories don't show, on average, a particular "trend", it behaves as a submartingale if, on average, the "trend" is increasing, and as a supermartingale if, on average, the "trend" is decreasing. (Borrelli, 2012). We will now begin our formal study of martingales:

$$E[M(t) \mid \mathcal{F}_t] = S(t)$$

For all $0 \leq s < t \leq T$.

We claim that a martingale defines what we call “a fair game”, then the conditional expectation of the future payoff $X$ at time $t$ is exactly its current price:

$$E[S(T)|\mathcal{F}_t] \; = \; S(t)$$

As we stated before, by intuition, a martingale is a process that is “constant on average”. Given all information up to time $s$, the best guess for the value of the process at time $t \geq s$. is the current value. Likewise, the concept of supermartingale and submartingale are similar and can be considered as an extension of this concept. We can give a mathematical definition.

1. A submartingale has the form

$$E[X_{n+1} \mid X_1, \ldots, X_n] \geq X_n$$

2. A supermartingale has the form

$$E[X_{n+1} \mid X_1, \ldots, X_n] \leq X_n$$

(Borrelli, 2012).

## 2.8 Markov process

In order to determine the stocks returns, Markov process is a helpful stochastic process.

Markov processes have no memory. Roughly speaking, it means that the prediction of the future is independent of its past, given its present value. Consequently, the variable’s past history and also the path through which a certain situation has been generated is not relevant. (Hull, 2002).

Stock prices in general are assumed to follow a Markov process, for example if the current price of a stock $S$t = 100, this is the only relevant piece of information.

An adapted process, $S$, is a Markov process if for any $t$, δ$t$ we have:

$$E[S_{t+\Delta t}|\, \sigma(S_s : \; s \; \leq \; t)] \; \equiv \; f(S_t)$$

for an appropriate deterministic function $f$.

This definition makes clearer what we stated above showing how the expected value of our process $St$, the price of the stock, after an increase $\Delta t$, is dependent only on the current time, $t$. In the definition above, past information is indicated with the expression σ, which represents the entire history of the process $S$ up to the point of time $s$. A specific type of Markov stochastic process in a continuous time setting is the Wiener Process.

## 2.9 Brownian Motion

The term Brownian derives from the name of the botanist Robert Brown that in 1827 observed that the movement of pollen grain suspended in a liquid (e.g. water) follows chaotic and irregular movements. The particle is much bigger than liquid molecules so we can see only the particle. As the molecules are moving, they hit to the particle producing a very small energy, but together it can move with the particle. Since we are not able to calculate motion equation for every molecule, we cannot say where will the particle move, but we can for example know what the probability is to find a particle in some area. Albert Einstein in 1905 formulated a mathematic model of Brownian Motion. But already in 1900 L.Bachelier had used Brownian Motion to describe the movement of stock prices and other financial indices on the Paris stock market (Borrelli, 2012). The random motion of a particle in a fluid subject to collision and influence of other particle is called Brownian Motion. One of the mathematical models of this motion is the Wiener Process (Krishnan, 2006).

In mathematical terms we say that in a filtered space $(\Omega, (\mathcal{F}_t)_{t\in[0,T],}\mathbb{P})$, taken $t$ as a continuous time parameter, $W = (W(T))_{t\leq T}$ is a Brownian motion if:

- W(0) = 0
- W is adapted to the filtration
- for any $s<t$, the increment W($t$)−W($s$) is independent of $\mathscr{F}_s$, and has distribution $N(0, t-s)$
- W(t) has continuous sample paths

From this definition we can say that:

- marginal distributions are Gaussian, for any t we can write W(t)−W(0) which is normally distributed with N(0,t)
- Brownian Motions have increments identically independent: for any u < s we can conclude that $W(u), W(t) - W(s)$ are independent and therefore have a joint normal distribution
$$N\left(\begin{pmatrix}0\\0\end{pmatrix}, \begin{pmatrix}u & 0\\0 & t-s\end{pmatrix}\right)$$

In general, fixing $0 \leq t_1 < t_2 < \ldots < t_n \leq T$ we obtain $n$ increments $W(t_1), W(t_2) - W(t_1), \ldots, W(t_n) - W(t_{n-1})$ independent and jointly Gaussian distributed, with:

$$N\left(\begin{pmatrix}0\\ \vdots\\ 0\end{pmatrix},\begin{pmatrix}t_1 & 0 & \dots & \dots & 0\\ 0 & t_2-t_1 & 0 & \dots & 0\\ \vdots & \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & 0 & \dots & t_n-t_{n-1}\end{pmatrix}\right)$$

In addition to being Markov processes, Brownian motion are also martingales. We prove it by considering two different dates $s < t$ and write $W(t) = W(t) - W(s) + W(s)$. The conditional expectation is:

$$E[W(t)\,|\mathcal{F}_s] = E[W(t) - W(s) + W(s)\,|\mathcal{F}_s] = W(s) + E[W(t) - W(s)] = W(s)$$

We now prove that a Brownian Motion is a martingale. Fixed two dates s < t we can write W(t) as W(t)−W(s)+W(s). We use this particular expedient to reach the conclusion that:

$$E[W(t)|Fs] = E[W(t) - W(s) + W(s)|Fs] = W(s) + E[W(t) - W(s)] = W(s)$$

## 2.10 Linear and Geometric Brownian motion

A linear transform of W, the standard Brownian motion, is:

$$B(t) = \mu t + \sigma W(t)$$

This is also referred to as Brownian Motion with drift.

**Geometric Brownian Motion**

Just as Brownian Motion is a Markov process, so is geometric Brownian motion, so the future given the present state is independent of the past.

Let’s take a Brownian Motion with drift μ and volatility σ which describes the path of a process Y.

The exponential transform is:

$$Y\,(t) = exp(X(t)) = exp(\mu t + \sigma W(t))$$

This expression is called Geometric Brownian Motion and, since it is the exponential of a Gaussian variable, it has a lognormal distribution.

Then we demonstrate that Geometric Brownian Motion are also Martingales: if we write $W(t) = W(t) - W(s) + W(s),\ for\ s < t$ we get the expectation:

$$E[exp(\mu t + \sigma W(t))|\mathcal{F}_s] = exp(\mu t + \sigma W(s))E[exp(\sigma(W(t) - W(s))]$$

and the expectation is a Gaussian variable and has a normal distribution: $N \sim (0, t-s)$, therefore is equal to $exp(\frac{\sigma^2}{2} (t-s))$ , in fact we obtain:

$$E[exp(\mu t + \sigma W(t))|\mathcal{F}_s] = exp(\mu t + \frac{\sigma^2}{2}(t-s) + \sigma W(s))$$

We can notice that the Geometric Brownian Motion is a Martingale if and only if the drift

$\mu = -\frac{\sigma^2}{2}$ because:

$$E[e^{\sigma(\sqrt{t}-s)x}]$$

Where $x$ is an extraction from a random normal distribution and:

$$\int e^{\sigma(\sqrt{t-s})x} \frac{e^{-\frac{1}{2}x^2}}{\sqrt{2\Pi}} dx$$

$$\int \frac{e^{-\frac{1}{2}\left(x^2 - 2\sigma(\sqrt{t-s}) + \left(\sigma(\sqrt{t-s})\right)^2\right) + \frac{\left(\sigma(\sqrt{t-s})\right)^2}{2}}}{\sqrt{2\Pi}} dx$$

$$e^{\frac{\left(\sigma(\sqrt{t-s})\right)^2}{2}} \int_{-\infty}^{+\infty} \frac{e^{-\frac{1}{2}\left(x - \sigma(\sqrt{t-s})\right)^2}}{\sqrt{2\Pi}} d(x - \sigma(\sqrt{t-s}))$$

given that the argument of the integral is the density function of a normal variable

$(x - \sigma(\sqrt{t} - s))$ it simplies to 1 given that the interval of validity is $[-\infty, +\infty]$ and we get that:

$$E\left[e^{\sigma(\sqrt{t-s})x}\right] = e^{\frac{\left(\sigma(\sqrt{t-s})\right)^2}{2}}$$

So, it is demonstrated that the Geometric Brownian Motion is a Martingale since all term depending on $t$ disappears.

## 2.11 Ito process

An Ito process is another type of stochastic process, it is a generalized Wiener process in which parameters $a$ and $b$ are functions of the underlying value of the variable $x$ and of the time variable $t$; the process is expressed with the equation:

$$dx_t = a(x_t, t)dt + b(x_t, t)dz_t$$

where $a$ and $b$ are respectively drift and standard deviation.

Either the expected drift rate and the variance rate of an Ito process can change value over time. The Ito's Lemma, named after the mathematician K.Ito, is an important process in the understanding of the behavior of functions of stochastic variable. (Borrelli,2012). Assuming that a variable $x$ follows the Ito process:

$$dx = a(x_t, t)dt + b(x_t, t)dz$$

Ito's lemma shows that a function F of $x$ and $t$ follows the process according to the equation:

$$dF(t,x) = F_t(t,x)dt + F_x(t,x)dx$$

Using the Taylor expansion, we make a second-order approximation:

$$dF(t,x) = F_t(t,x)dt + F_x(t,x)dx + \frac{1}{2}(F_{xx}(t,x)(dx)^2 + 2F_{tx}(t,x)dtdx + F_{tt}(t,x)(dt)^2)$$

Usually second order elements of an approximation are negligible and are rarely considered. We now take $t$ as the time parameter and consider a function $Y$ which depends on time and on a Brownian motion $W$. We consider the following:

$$Y(t) = F(t, W(t))$$

Using our second order approximation we consider the following variation for $Y$ in terms of $t$ and W:

$$Ft(t,W(t))dt + Fx(t,W(t))dW(t) + \frac{1}{2}(Fxx(t,W)(dW(t))^2 + 2Ftx(t,W(t))dtdW(t) + Ftt(t,W(t))(dt)^2)$$

Following the intuition that $dW(t) = W(t+dt) - W(t) \sim N(0,dt)$, we can approximate the square increment $(W(t))^2$ with its mean:

$$(W(t))^2 \sim dt$$

Ito's Lemma sums up what our findings where so far, and stated the following:

Let $F(t, x)$ be a smooth function. The Markov process defined by:

$$\mathrm{F(t, W(t))}$$

has dynamics given by the following stochastic differential equation:

$$dF\big(t, W(T)\big) = \left(F_t\big(t, W(t)\big) + \frac{1}{2}F_{xx}\big(t, W(t)\big)\right)dt + F_x(t, W(t)dW(t)$$

Another name given to an Ito process, diffusion process, is any adapted process Y whose dynamics may be written as:

$$dY\ (t)\ =\ \alpha(t)dt\ +\ \beta(t)dW(t)$$

where α and β are two coefficients. The first one, α, is referred to as the drift of the process. In reality though, in Finance the practice is to call drift the fraction $\frac{\alpha(t)}{Y(t)}$ . The second coefficient, β, is the diffusion of the process. In the case of Brownian motions with drift and Geometric Brownian motions (S) we have the following conditions. B verifies:

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

while the Geometric Brownian motion, $S(t)\ =\ exp(bt\ +\ \sigma dW(t))$ satisfies:

$$dS(t)\ = \left(b\ +\ \frac{\sigma^2}{2}\right)S(t)dt\ +\ \sigma S(t)dW(t)$$

where sometimes we can call $\mu\ =\ (b\ + \frac{\sigma^2}{2}\ )$.

## 2.12 Black-Scholes

Here we want to present Black-Scholes model, usually for pricing derivative. In order to derive some solution, we have to simplify the model that we would be able to formulate with mathematical structure we already know. Let's say we have only two assets, a risk-free asset (B, the "bond") and a risky stock. The former pays continuously an interest $r \geq 0$ and has the following characteristics:

$$\begin{cases} dB(t) = rB(t)dt \\ \quad B(0) = 1 \end{cases}$$

Or $B(t) \geq e^{rt}$ .

The latter on the other hand satisfies a stochastic differential equation with an initial condition (Cauchy's Problem). It is described as:

$$\begin{cases} dS(t) = \mu S(t)dt + \sigma S(t)dW(t) \\ S(0) = S_0 \end{cases}$$

here, $S_0$ is the observed, current, market price of the risky stock and the terms μ and σ are constants, with $\sigma \geq 0$, respectively called drift and volatility.

Solving the Cauchy problem and finding its unique solution, leads us to find out that $S$ must satisfy:

$$S(t) = S_0 e^{\left(\mu - \frac{\sigma^2}{2}\right)t + \sigma W(t)}$$

because of the lognormal property of the Geometric Brownian motion, we state that the marginals of the process $S(t)$ satisfy:

$$ln \frac{S(t)}{S_0} \sim N\left(\left(\mu - \frac{\sigma^2}{2}\right)t, \sigma^2\right)$$

meaning that the mean and the variance of the stock logreturns grow over time linearly. Given these properties m is defined as the exponential growth of the average stock price, while volatility is the standard deviation of the annual log return.

## 2.13 Lognormal property

If we take a function $Y$ of $S$ such that:

$$Y = ln\, S$$

practically we want to derive the process of the $ln\, S$, where $dS = \mu dt + \sigma dW(t)$.

Using Ito's Lemma we can state that the process followed by $Y$ is:

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

Because:

$\frac{\delta G}{\delta S} = \frac{1}{S} \qquad \frac{\delta^2 G}{\delta S^2} = -\frac{1}{S^2} \qquad \frac{\delta G}{\delta t} = 0$

It implies that:

$$lnS_t - lnS_0 \sim N\left(\left(\mu - \frac{\sigma^2}{2}\right)\right)T, \sigma^2 T$$

And

$$lnS_T \sim N\left(lnS_0 + \left(\mu - \frac{\sigma^2}{2}\right)T, \sigma^2 T\right)$$

Since $\ln\big(S(t)\big)$ has normal distribution and $S(T)$ is log-normal, it holds that the stock price at time T has a log-normal distribution.

# 3. Portfolio Optimization

## 3.1 Portfolio theory

Portfolio optimization is a major area in finance. We know that a basic premise of economics is that, due to the scarcity of resources, all economics decisions are made in the face of trade off. The trade-off facing the investors are risk versus expected return. The investment decisions are not merely which securities to own, but how to divide investor's wealth among securities, given that the main objective is to maximize the yield and at the same time minimize the risk. This is the problem of "Portfolio Optimization". The publication of Harry Markowitz's landmark paper, "Portfolio Selection", in 1952 in the Journal of Finance, can be considered the moment of the birth of modern financial economics (Rubinstein, 2002). He suggested one way to optimizing a portfolio and in 1990 his work earned him a share of 1990 Nobel Prize in Economic Science for his pioneering theoretical contribution to portfolio theory, together with Merton Miller and William Sharp (Amu, Millegard , 2009). Though widely applicable, mean-variance optimization has had the most influence in the practice of portfolio management. In his work Markowitz poses the underpinning for Modern Portfolio Theory, providing a framework to construct and select an investment portfolio, that, as we stated before, is based on the maximization of expected returns and simultaneous minimization of investment risk. The foundation for this theory was then expanded by William Sharpe, awarded of the Nobel Prize, known for his Capital Asset Pricing Model on the theory of financial asset price formation (Mangram, 2013).

As a matter of fact, after the initial selection procedure, the agent become a passive agent as he can only be a spectator of price fluctuations. It is however relevant to introduce Markowitz model before the discussion of a more advance intertemporal model.

## 3.2 Markowitz portfolio

The primary goal of the Modern Portfolio Theory established by Markowitz in his article "Portfolio Selection" (Markowitz, 1952) was the description of the impact of asset allocation that is commonly known as portfolio diversification by the number of securities within the portfolio and their covariance relationship. The ideas introduced in this article have come to build the basis of what is now called mean-variance analysis and Modern Portfolio Theory. (Fabozzi et al., 2007)

Markowitz developed mean-variance analysis in the context of selecting a portfolio of common stock. Mean-variance analysis requires knowledge of expected return and standard deviation on each asset, as well as the correlation of returns for each pair of assets because of the opportunity to reduce total portfolio risk comes from the lack of correlation across assets. We will describe more in detail Markowitz Model.

First, as with any model, we have to underline the assumptions on which the model is based:

*"1. Investors are rational, so they seek to maximize the expected return while minimizing risk.*

*2. All investors have the same expected single period investment horizon.*

*3. All investors are risk-adverse, that is they are only willing to accept greater risk if they are compensated with a higher expected return.*

*4. Investors base their investment decisions on the expected return and risk, so they receive all pertinent information.*

*5. Markets are perfectly efficient."* (Mangram, 2013).

Mean- variance analysis is based on a single period model of investment. This means that the investor, at the beginning of the period is active, and he decides how to allocate his wealth among various assets, during the period he is an observer and each asset generates a rate of return, at the end of the period, final time T, his wealth has been changed and the investor collect the results of his investment. (Rubinstein, 2002).

We consider a market with $d$ securities with prices $p1, p2, .., pd > 0$ at time 0 and final prices $P1, P2, .., Pd$ at the final time T. the securities are not feasible. They are random variables over the probability space ($\Omega$, $\mathscr{F}$, $\mathbb{P}$ ), the return is:

$$Ri(T) = \frac{Pi(T)}{p_i}, with\ i = 1, .., d.$$

Then we assume that the mean and the covariance are given, where: $E[R_i(T)] = \mu_i$ and $cov(Ri(T), R_j(T) = \sigma_{ij}$.

Assuming that each security is perfectly divisible, that is we can hold $\varphi_i \epsilon R$ of security $i$, $i$ = 1,….,$d$.

An investor with an initial wealth $X$ >0 is assumed to hold $wi \geq 0$ shares of security $i$ with

$$\sum_{i=1}^{d} w_i P_i = X$$

.

Then the portfolio vector is $\pi = (\pi 1, \pi 2, \ldots, \pi d)$ and the portfolio return $R^{\pi}$ are given as:

$$\pi i = \frac{wiPi}{X}$$

$$R^{\pi} = \sum_{i=1}^{d} \pi_i R_i(T)$$

The component of the portfolio vector represents the fraction of wealth which are invested in the corresponding assets of securities.

We can now write the mean and variance of the portfolio as:

$$E(R^{\pi}) = \sum_{i=1}^{d} w_i \mu_i$$

$$var(R^{\pi}) = \sum_{i=1}^{d} \sum_{j=1}^{d} \pi_i \sigma_{ij} \pi_i$$

As we stated above, Markowitz reasoned that investor should decide on the basis of a trade-off between risk and expected return. He suggested that risk should be measured by the variance of returns, while the return by the portfolio mean.

Following Markowitz, the investors' problem is a constrained minimization problem in the sense that an investor who seeks to minimize the variance of the portfolio given a lower bound $C_2$ must seek:

$$\min Var(R_i)$$

Subject to: $w_i \geq 0, \sum_{i=1}^{n} w_i = 1,\ E(R_i) \geq C_2, i = 1..n$

For an investor who seeks to maximize the portfolio return, given an upper bound $C_1$, the problem will be formulated as:

$$\max E[R_i]$$

Subject to the constraints: $w_i \geq 0, \sum_{i=1}^{n} w_i = 1\,, Var(R_i) \leq C_1, i = 1..n$

To solve these problems first order conditions are used.

## 3.3 Introduction to Merton Problem

A more advanced approach to portfolio selection which accommodates for the risk aversion of the investor and is not subject to the static nature of the Markowitz mean-variance approach was presented by Robert Merton in 1969. In order to have a clear environment, first we introduce the CRRA utility function.

## 3.4 Preferences and risk aversion

We introduce the form and discuss the properties of the utility functions that investors are assumed to possess in this model. They play a key role in determining how to maximize the outcome by means of our strategy. Merton in the formulation of his theory takes into consideration the CRRA utility function, that is the constant relative risk aversion function. CRRA are a specific set of utility function which belongs to the family of HARA. The idea is that the HARA form is the only form of function able to satisfy some optimization related economic principles. (Perets et al. 2013). The only functions of this wider set of functions that satisfy the CRRA property are referred to as isoelastic functions and are of the form:

$$u(t,x) = e^{-\rho t}\frac{x^{1-R}}{1-R}$$

Or:

$$u(t,x) = -\gamma^{-1}\exp(-\rho t - \gamma x)$$

Where $\rho, \gamma, R > 0$ $and$ $R \neq 1$ and R represents the risk aversion of the consumer.

In analyzing Merton Problem we will make use of the first type of utility.

The use of a CRRA utility is convenient as we can take advantage of the scaling properties as we will see after.

## 3.5 Merton Problem

Merton considered a situation in which the investor had the limited choice of investing his wealth in only two different assets: a risky asset and a risk-free asset. We have to underline that, given a limited time horizon, the investor's objective, who is risk averse, was to maximize the expected utility of his wealth at the end of the time range taken into account. Merton's goal was to determine how the investor should allocate and reallocate his wealth at each time point in order to achieve the previously selected goal. (Merton, 1992).

We characterize the dynamics of the agent's wealth trough the equation:

$$dw_t = r_t w_t dt + n_t(dS_t - r_t S_t dt + \delta_t dt) + e_t dt - c_t dt)$$

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

For some given initial wealth $w_0$. Here *e* and *c* are respectively an endowment stream and a consumption stream, the process r is an adapted process, interpreted as the riskless rate of interest. D, S, r and *e* will generally be assumed given, as the initial wealth $w_0$. The agent can control the portfolio process *n* and the consumption process c. The model allows for changes in the quantity of wealth consumed and in the number of stocks purchased at time *t*. The investor initially chooses how much of $w_0$ to invest in stocks and how much in the risk-free asset. We can think of the latter as wealth in bank account that returns an interest r, after setting such values, the total worth of the stocks bought will be $n_0 S_0$ and the bank account will add up to $w_0 - n_0 S_0$. Because the values chosen for c and n can change over time, due to portfolio adjustments, it is important to specify that the pair $(n, c)$ which makes up the investor's strategy on how much to invest and to consume, has to be admissible at any time. (Rogers, 2013).

The definition for this condition if the following: a pair $(n_t, c_t)_{t\geq 0}$ is said to be admissible for initial $w_0$ if the above process stays non-negative at all times. We will write the set of all admissible strategies as:

$$\mathcal{A} \equiv \cup_{w>0} \mathcal{A}(w)$$

The portfolio of an investor can be characterized by the number of units of the asset held, or by the cash values invested in the different assets. Depending on the particular context, either one may be preferable. As a convention we say that *n* stays for the number of assets and θ for what the holding of assets is worth. Thus, if at time t we hold $n_t^i$ units of asset i, whose time t-price is $S_t^i$, then we have the identity:

$$\theta_t^i = n_t^i S_t^i$$

we can also characterize the portfolio held in terms of the proportion of wealth assigned to each of the assets. And so, given that $\pi_t^i$ is the proportion of wealth invested in asset *i* at time *t*, we will write the notations as:

$$\theta_t^i = n_t^i S_t^i = \pi_t^i w_t$$

We now suppose that the problem for the agent is to choose the pair $(n, c)$ such that:

$$\sup_{(n,c)\in\mathcal{A}(w_0)} E\left[\int_0^T u(t,c_t)dt + u(T,w_T)\right]$$

Where the function $u$ is assumed to be concave increasing in its second argument, and measurable in its first. The time horizon T is taken to be a positive constant. The investor aims to maximize the final expected value of the intertemporal and final utilities over a specific time horizon.

We can include some special cases:

-The infinite-horizon problem:

$$\sup_{(n,c)\in\mathcal{A}(w_0)} E\left[\int_0^\infty u(t,c_t)dt\right]$$

And the terminal wealth problem:

$$\sup_{(n,0)\in\mathcal{A}(w_0)} E[u(w_T)]$$

We now introduce the Davis-Varaija Martingale Principle of Optimal Control to solve the agent's problem of achieving his objective such that the wealth process $w$ generated by his dynamics remains non-negative. We can use several methods, but an important principle underlying many of the approaches is the Martingale Principle of Optimal Control. It states:

Suppose we have an objective of the form

$$\sup_{(n,c)\in\mathcal{A}} E\left[\int_0^T u(t,c_t)dt + u(T,w_T)\right]$$

And there exists a function $V : [0,T]\times\mathbb{R}^+ \to \mathbb{R}$ which is continuous in at least its second derivative and such that $V(T,\cdot) = u(T,\cdot)$. Suppose also that for any $(n,c) \in \mathcal{A}(w_0)$:

$$Y_t \equiv V(t,w_t) + \int_0^t u(s,c_s)ds \text{ is a supermartingale}$$

Then for some $(n^*,c^*) \in \mathcal{A}$, the process Y is a martingale. Then such pair $(n^*,c^*)$ is optimal and the value of the problem starting from $w_0$ is therefore:

$$V(0,w_0) = \sup_{(n,c)\in\mathcal{A}(w_0)} E\left[\int_0^T u(t,c_t)dt + u(T,w_T)\right]$$

We prove it from the supermartingale property of Y. For any $(n, c) \in \mathcal{A}(w_0)$:

$$Y = V(0, w_0) \geq E[Y_T] = E\left[\int_0^T u(t, c_t)dt + u(T, w_T)\right]$$

Holds. Making use of $V(T,\cdot) = u(T,\cdot)$ , so the value is not greater than $V(0, W_0)$ and in the case of $(n^*, c^*)$ the inequality becomes an equality, therefore $(n^*, c^*)$ is optimal. (Rogers, 2013).

## 3.6 The value function approach

The value function approach is the classical methodology for solving a stochastic optimal control problem, based on the Martingale Principle of Optimal Control. We suppose that our asset follows a Geometric Brownian Motion and its dynamics are the following:

$$dS_t^i = S_t^i\left(\sum_{j=1}^{N} \sigma^{ij} dW_t^j + \mu^i dt\right)$$

Where $\sigma^{ij}$ and $\mu^i$ are constants and W is a d-dimensional Brownian Motion and our risk-less rate *r* is constant and that the endowment *e* and the dividend payout δ are both equal to zero. To make it more compact we re-write the equation as:

$$S_t = S_t(\sigma \cdot dW + \mu dt)$$

And now we can express the wealth dynamics in terms of θ (recall that $\theta_t^i = n_t^i S_t^i$):

$$dw_t = rw_t dt + \theta_t \cdot (\sigma dW_t + (\mu - r)dt) - c_t dt$$

Given that we want to find a function that satisfies the MPOC condition we proceed by writing down our process Y and perform an Ito expansion, assuming that V is sufficiently regular:

$$dY_t \;=\; V_t dt \;+\; V_w dw \;+ \frac{1}{2}\, V_{ww}(dw)^2 \;+\; u(t, c)dt$$

$$= dY_t \;=\; V_w \theta \cdot \sigma dW \;+ \left\{u(t, c) \;+\; V_t \;+\; V_w(rw \;+\; \theta \cdot (\mu - r) - c) \;+\; \frac{1}{2}|\sigma^T \theta|^2 V_{ww}\right\} dt$$

the condition for the Ito expansion to be a supermartingale is that the drift be non-positive, where the drift is equal to zero, the Ito expansion is a martingale. This happens at the optimal strategy ($\theta^*$,$c^*$), at this point the funtion V is our value function. So we consider:

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

This (non-linear) partial differential equation (PDE) for the unknown value function V is called Hamilton-Jacobi-Bellman (HJB) equation and it will be fundamental for finding a solution. Anyway, we have to deal with some questions: Is there any solution to the partial differential equation? If so, is there a unique solution that satisfy the boundary conditions? Is the supremum attained? Is V actually the value function?

To answer these questions HJB is a good way, in fact if we can find a function V which solves the HJB equation then, by direct means, it will be possible to verify that this function V is the value function. (Rogers,2013)

In any case, also if is possible to answer the questions above, we just know that there is a value function and that it is the unique solution, we miss some information: how the optimal value ($\theta^*$,$c^*$) will looks or how the solutions will change when we change any of the input parameters. Usually in order to find a more explicit solution we have to assume a simple form of utility that is:

$$u(t,x) \equiv e^{-\rho t} u(x) \equiv e^{-\rho t} \frac{x^{1-R}}{1-R}$$

(Rogers, 2013).

## 3.7 Infinite horizon Merton Problem

We now analyze the infinite-horizon Merton Problem. We focus on and we assume the case of CRRA utility (with the properties we stated above) form for $u$, so we write:

$$u(t,x) \equiv e^{-\rho t} u(x) \equiv e^{-\rho t} \frac{x^{1-R}}{1-R}$$

Thus, we wish to solve for

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

And the admissible $(n, c)$ which attains the supremum.

This problem can be solved using a specific methodology:

1. Make a guess on the form of the solution

2. Make use of HJB equation to find the solution

3. Find a simple bound for the value of the underlined problem

4. Verify that for the optimal solution the bound is attained

STEP 1:

We can write down the form of the solution easily because of the scaling property, by linearity of our wealth equation (we will explain this right after) and so we can now state that:

$$V(w) = \gamma_M^{-R} u(w) \equiv \gamma_M^{-R} \frac{w^{1-R}}{1-R}$$

For some constant $\gamma_M > 0$. The problem is reduced as the finding of such $\gamma_M$..

By linearity of wealth equation, it for any $k > 0$

$$V(kw) = \sup_{(n,c)\in\mathcal{A}(kw)} E\left[\int_0^\infty e^{-\rho t} u(c_t)dt\right]$$

$$= \sup_{(n,c)\in\mathcal{A}(kw)} E\left[\int_0^\infty e^{-\rho t} u(kc_t)dt\right]$$

$$= \sup_{(n,c)\in\mathcal{A}(kw)} k^{1-R} E\left[\int_0^\infty e^{-\rho t} u(c_t)dt\right]$$

$$= k^{1-R} V(w)$$

STEP 2:

Using the HJB equation we can identify $\gamma_M$ . To find the unknown constant $\gamma_M$ we take:

$$V\,(t,w) \;=\; \sup_{(n,c)\in\mathcal{A}(w)} E\left[\int_0^\infty e^{-\rho t} \frac{{c_s}^{1-R}}{1-R} ds \mid w_t = w\right]$$

because of the time-homogeneity of the problem:

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

Where V is defined as:

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

and because of the scaling form of the solution:

$$V(w) = \gamma_M^{-R} u(w) \equiv \gamma_M^{-R} \frac{w^{1-R}}{1-R}$$

we conclude that:

$$V(t,w) = e^{-\rho t} \gamma_M^{-R} u(w)$$

Now that we have the form of the solution, we just must identify the value of the constant. We make use of the HJB equation that involves optimize over the two factors that make up the investment strategy, θ and c.

Optimization over θ:

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

From here $\theta^*$ :

$$\theta^* = -\frac{V_w}{V_w w}(\sigma\sigma^T)^{-1}(\mu - r)$$

Using $V(t,w) = e^{-\rho t}\gamma_M^{-R} u(w)$ we end up with the optimal wealth allocation in each asset *i*:

$$\theta^* \ = \ wR^{-1}\,(\sigma\sigma^T)^{-1}(\mu - r)$$

and now we can introduce the notation:

$$\pi_M \ \equiv \ R^{-1}\,(\sigma\sigma^T)^{-1}(\mu - r)$$

a constant N-vector as the Merton Portfolio.

$\theta^* \ = \ wR^{-1}\,(\sigma\sigma^T)^{-1}(\mu - r)$ tells us that for each asset *i* and for all time *t* > 0 the cash value of the optimal holding of asset *i* will be proportional to the current wealth $w_t$ with $\pi_M^i$ as constant of proportionality. So we will have:

$$(\theta_t^*)i \ = \ w_t \pi_M^i$$

Optimization over c is achieved through the introduction of the following convex dual function of u:

~

$$\tilde{u}(y) \ \equiv \ sup\{u(x) - xy\}$$

Then we have for $u(x) \ = \ \frac{x^{1-R}}{1-R}$ that:

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

where $\tilde{R} = R^{-1}$.So the optimization over c becomes as:

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

As we did for θ we now substitute $V(t,w) = e^{-\rho t}\gamma_M^{-R}u(w)$ and we obtain:

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

Again, what we obtain is an optimizing $c^*$ proportional to *w*:

$$c^* = \gamma_M w$$

This does not come as a surprise because of the scaling property of the objective.

We now put all together and return the candidate value function $V(t,w) = e^{-\rho t}\gamma_M^{-R}u(w)$ to the Hamilton-Jacobi-Bellman equation as a result we have:

$$0 = e^{-\rho t}[\frac{R}{1-R}(\gamma_M w)^{1-R} - \rho\gamma_M^{-R}u(w) + rw\gamma_M^{-R}w^{-R} + \frac{1}{2}\gamma_M^{-R}w^{1-R}\frac{|\kappa|^2}{R}$$

$$= \frac{e^{-\rho t w^{1-R}}\gamma_M^{-R}}{1-R}[R\gamma_M - \rho - (R-1)(r + \frac{1}{2}\frac{|\kappa|^2}{R}]$$

where $\kappa \equiv \sigma^{-1}(\mu - r)$ is the market price of risk vector. From this the value of $\gamma_M$ is:

$$\gamma_M = R^{-1}\left\{\rho + (R-1)(r + \frac{1}{2}\frac{|\kappa|^2}{R}\right\}$$

The value function of the Merton Problem then is $VM(w) \equiv V(t,w)$,

$$V_M(w) = \gamma_M^{-R}u(w)$$

We now can write down the form of the optimal solution over the infinite-horizon problem thus, we believe that, we invest proportionally to wealth and we also consume proportionally to wealth. The constants of proportionality are given by the $\theta^*$ and $c^*$ we computed.

To conclude two issues must be taken into account:

What happens if the expression $\gamma_M = R^{-1}\left\{\rho + (R-1)(r + \frac{1}{2}\frac{|\kappa|^2}{R}\right\}$ for $\gamma_M$ is negative? And can we prove that $\gamma_M$ is actually the optimal solution?

The first question relates directly to whether the Merton Problem is well-posed or not.

The second one is more general for this reason  we will first answer to this one, to answer we first should make the assumption that $\gamma_M > 0$, where $\gamma_M$ is given by $\gamma_M \; = \; R^{-1}\left\{\rho \; + \; (R-1)(r \; + \; \frac{1}{2}\frac{|\kappa|^2}{R}\right\}$.

Suppose we are given the initial $w_0$ we consider the wealth evolution $w^*$ under the conjectured optimal control. What we see is:

$$dw^*t \; = \; w_t^*\{\pi_M \; \cdot \sigma dW_t \; + \; (r \; + \; \pi_M \; \cdot (\mu - r) - \gamma_M)dt\}$$

$$= \; w_t^*\{R^{-1}\kappa \cdot dW_t \; + \; (r \; + \; R^{-1}|\kappa|^2 \; - \gamma_M)dt\}$$

which is solved by:

$$w_t^* = \; w_0 exp[R^{-1}\kappa \cdot W_t \; + \; (r \; + \frac{1}{2} \; R^{-2}|\kappa|^2(2R-1) - \gamma_M)t]$$

STEP 3:

The proof of optimality is based on a trivial inequality:

$$u(y) \leq u(x) + (y - x)u'(x)$$

for $(x, y) > 0$ which express the fact that the tangent to the concave function $u$ at $x > 0$ lies everywhere above the graph of $u$. if we consider any admissible $(n, c)$ then we can bound the objective to

$$E\left[\int_0^\infty e^{-\rho t}\, u(c_t)dt\right] \leq E\; \left[\int_0^\infty e^{-\rho t}\,\{u(c_t^*) + (c_t - c_t^*)u'(\, c_t^*)dt\right]$$

$$= E\left[\int_0^\infty e^{-\rho t} u(c_t^*)dt\right] + E\left[\int_0^\infty (c_t - c_t^*)\zeta_t dt\right]$$

That we can simplified as:

$$\zeta_t \equiv e^{-\rho t}u(c_t^*) \propto \exp(-k \cdot W_t - (r + \frac{1}{2} \mid k \mid^2 \, t)$$

Now $\zeta_t$ is the state-price density, also named stochastic discount factor. A property related is that for any admissible $(n, c)$:

$$Y_T \; \equiv \; \zeta_t w_t \; + \int \zeta_s c_s ds$$

is a local martingale.

This is provable through Ito's calculus from the wealth equation. And from here we can be sure, given that wealth and consumption are non-negative, that Y is a non-negative supermartingale and thus:

$$w_0 = Y_0 \geq E[Y_\infty] \geq E\left[\int_0^\infty \zeta_s c_s ds\right]$$

STEP 4:

Last step is to verify that:

$$w_0 = E\left[\int_0^\infty \zeta_s c_s^* ds\right]$$

That is, verifying that the optimal supreme is attained. Here $c^*$ is the optimal consumption process. We combine the bound condition and the conditions for $w_0$ so finally we can write down that for ay admissible $(n, c)$:

$$E\left[\int_0^\infty e^{-\rho t} u(c_t) dt\right] \leq E\left[\int_0^\infty e^{-\rho t} u(c_t^*) dt\right]$$

Which verifies that $(n^*, c^*)$ is the optimal solution. (Rogers, 2013)

## 3.8 Finite horizon Merton Problem

The same technique we used for the infinite horizon can be used here, so the solution to the finite horizon Merton problem will be like the infinite horizon case. The first that constant relative risk aversion (CRRA) holds in consumption and that the utility function *u* is separable. The agent's goal is therefore:

$$\sup E\left[\int_0^T h(t) u(c_t) dt + A u(w_t)\right]$$

for strictly positive function $h$ and constant $A > 0$. Moreover, $u'(x) = x^{-R}$ for some $R > 0,\ R \neq 1$. We can again take advantage of the scaling properties of the CRRA functions and get that the value function:

$$V(t, w) = \sup\left[\int_t^T h(t) u(c_t) dt + A u(w_T) \mid w_t = w\right]$$

which for some function *f*, must have the form:

$$V\,(t, w) = f(t) u(w)$$

The Hamilton-Jacobi-Bellman equation for this problem is:

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

If we substitute our scaled form $V(t,w) = f(t)u(w)$ of the function into the previous equation, we obtain:

$$0 = \sup_{y,q} u(w)[f' + (r + y(\mu - r) - q)(1 - R)f - \frac{1}{2}R(1 - R)\sigma^2 y^2 f + hq^{1-R}]$$

where $y = \theta/w$ and $q = c/w$. The conditions for optimality are now given by:

$$y = \pi_M \ \ and \ \ f = hq^{-R}$$

We can then conclude that investment is actually the same, but in general now we are not consuming at a constant rate multiple of the initial wealth. So in conclusion:

$$\theta_t^* = \pi_M w_t \ and \ c_t^* = w_t \left(\frac{h(t)}{f(t)}\right)^{1/R}$$

Comparing to the infinite horizon case, we have the same investment strategy but now we no longer consume at a constant rate proportional to initial wealth. (Rogers, 2013)

# 4 Robust Merton

## 4.1 Introduction

P.Huber presented the idea of robustness with regards to statistical estimation of an unknown parameter. Huber presented the so-called gross error and demonstrated that an optimal estimate is a maximum likelihood estimate developed for the least favourable distribution. What does this analytically mean? it implies that we have to solve a minimax problem. In mathematical finance, most approaches and methodologies implicitly assume that the underlying asset model is completely determined: the parameters, trend and volatility, of the model are known. (Tevzadze et al, 2013.). The term “robust” is often interpreted to mean “minimax”, that is, the opponent chooses which probability model from a pre-specified set will be used in order to make your value as small as possible. (Rogers, 2013). In other words, when we are dealing with robust optimization, we have to take a worst case that is, as we stated above, a max-min approach. This means that we should find portfolio weights such that the return is maximized even when the asset return takes its worst value over the set, given that the uncertainty of the parameters varies in specific set, determined by the knowledge of the probability distribution. (Fabozzi et al., 2007).

Generally, financial modelling depends on the choice of an underlying P. P express the stochastic nature of market price evolutions, and the underlying risk factors, such as stock prices or interest rates, have been modeled as Markovian diffusion. Anyway, multifaceted nature of the global economic and financial dynamics renders imprecise the identification of P. So, we incorporate model uncertainty by replacing the single P by a set of probabilities $\mathcal{P}$, consisting of plausible models. (Biagini,Pinar, 2017).

The goal of the agent is maximize uncertainty averse utility from consumption and terminal wealth, in this case the investor has a pessimistic view of the odds, and takes a max-min way to deal with the problem, first minimizing a utility functional from wealth X over the plausible models $\mathcal{P}$ and then maximizes wealth X(θ) over portfolio strategies θ. We can see the robust problem as a zero-sum, two players game between the representative agent and the adverse market. A main distinction to be drawn in the literature on robust portfolio selection is the dominates vs nondominated family:

1) Case $\mathcal{P}$ is a dominated family. We say that family $\mathcal{P}$ is dominated if all $P \in \mathcal{P}$ are absolutely continuous with regard to a reference probability $P^0$, $P \ll P^0$.

We provide a definition of an absolutely continuous function:

**Definition 4.1** *A function* $f : [a,b] \to \mathbb{R}$ *is said to be absolutely continuous on* $[a,b]$ *if, given* $\varepsilon > 0$, *there exists some* $\delta > 0$ *such that*

$$\sum_{i=1}^{n} | f(y_i) - f(x_i) | < \varepsilon$$

*Whenever* $\{[x_i, y_i]: i = 1,..n\}$ *is a finite collection o mutually disjoint subintervals of* $[a,b]$ *with* $\sum_{i=1}^{n} | y_i - x_i | < \delta$.

*Clearly, an absolutely continuous function on* $[a,b]$ *is uniformly continuous on* $[a,b]$ *is uniformly continuous. Moreover, a Lipschitz continuous function on* $[a,b]$ *is absolutely continuous. Let* $f$ *and* $g$ *be two absolutely continuous functions* $[a,b]$. *Then* $f+g, f-g$ *and* $fg$ *are absolutely continuous on* $[a,b]$. *If, in addition, there exist a constant C > 0 such that* $| g(x) | \geq C$ *for all* $x \in [a,b]$, *then* $f/g$ *is absolutely continuous on* $[a,b]$.

*If* $f$ *is integrable on* $[a,b]$, *then the function F defined by*

$$F(x) \coloneqq \int_a^x f(t)dt, \qquad a \leq x \leq b$$

*Is absolutely continuous on* $[a,b]$.

A family $\mathcal{P}$ is dominated if all $P \in \mathcal{P}$ are absolutely continuous with regard to a reference probability $P^0$, $P \ll P^0$.

This happens when:

- $\Omega$ is discrete, $P^0 = \Sigma_{n \geq 1} \frac{1}{2^n} \delta_{\omega_n}$
- $\mathcal{P}$ is finite or countable; dominating $P^0 = \Sigma_n c_n P_n$
- In a diffusion context; when there is uncertainty only in the drift.

2) Case $\mathcal{P}$ is nondominated. This is the case when we estimate the volatility coefficient. Estimation comes with error intervals. We now specify the model in the non dominated case and the consumption/investment problem formulation. We assume that the asset prices process is a n-dimensional diffusion, wih a n-dimensional Wiener process. The agent is diffident about the constant drift and volatility estimates $\hat{\mu}$ and $\hat{\sigma}$. Thus she consider as plausible all the covariances matrix lying in a given compact set K, fairly general, satisfying a uniform ellipticity condition $\min_{\Sigma \in K} y'\Sigma\, y \geq h^2 \|y\|^2, h > 0$.

The agent then considers all the drift which take values in a ellipsoid centered at $\hat{\mu}$, shaped by $\Sigma$

$$U_\epsilon(\Sigma) = \{u \in \mathbb{R}^n \,|(u - \hat{\mu})'^{\Sigma^{-1}}(u - \hat{\mu}) \leq \epsilon^2 \}$$

In which $\epsilon \geq 0$ is the radius of ambiguity.

For a given path of $\sigma$, let $\Sigma_t(\omega) = \sigma_t(\omega)\sigma_t(\omega)'$. Then

$$\mu_t(\omega) \in U_\epsilon(\Sigma_t(\omega)) \;\forall t, \omega$$

## .4.2 The general Merton Problem under ambiguity averse agent

Let's consider an agent that invest in n risky assets and riskless asset. In particular we are dealing with the Black-Scholes-Merton market model assumptions, where r, our riskless rate, is constant and the n traded risky dynamics are:

$$dS_t^i = S_t^i \left( \sum_{j=1}^{N} \sigma^{ij} dW_t^j + \mu^i dt \right)$$

Where $\sigma^{ij}$ and $\mu^i$ are constants and W is a standard, *n*-dimensional Brownian motion on a filtered space $(\Omega, (\mathcal{F}_t)_{t \geq 0}, P)$. We now rewrite the equation in a matrix-vector form so:

$$dS_t = Diag(S_t)(\mu dt + \sigma dW_t)$$

Where $Diag(S_t)$ the diagonal $n\ x\ n$ matrix with *i-th* diagonal element equal to $S_t^i$,μ is a n-vector and σ is $n\ x\ n$

Matrix. Also, σ is required to be invertible, so that the instantaneous covariance matrix $\Sigma\ =\ \sigma\sigma'$ is also invertible. Given the initial endowment $x > 0$, the investor is allowed to trade and consume in a self-financing way. To be explicit, let $h\ =\ (h_t)_t$ denote the *n*-dimensional progressively measurable process, representing the number of shares of each asset held in portfolio, and let the progressively measurable, non-negative, scalar process c indicate the consumption stream. Assume also that $\int_0^{\cdot} h'_s \Sigma h_s\, ds$ and $\int_0^{\cdot} c_s\, ds$ are finite P–a.s. The wealth process X is governed by the following stochastic differential equation:

$$dX_t\ =\ (rX_t\ +\ h_t' Diag(S_t)(\mu\ -\ r\mathbf{1})\ -\ c_t)dt\ +h_t'\ Diag(S_t)\sigma dW_t$$

The pair $(\theta_t, c_t)$ if the solution to the above equation, which is defined P-a.s., remains P-a.s. non-negative at all times., given the initial wealth x. Let $\mathcal{A}^P(x)$ be the set of all admissible $(\theta, c)$ pairs for initial wealth x. Given a time horizon T, the agent is trying to choose $(\theta\ , c)\ \in\ \mathcal{A}^P(x)$ so as to maximize the expected utility from running consumption and terminal wealth:

$$\sup_{(\theta\ ,c)\in\mathcal{A}^P\ (x)} E\left[\int_0^T u(t, c_t)dt\ +\ u(T, X_T\ )\right]$$

This class of stochastic control problems is known under the name of Merton problem.

**The utility function:**

The utility function $u : (0,\infty) \times\ \mathbb{R}^n \rightarrow (-\infty,\infty)$ is jointly measurable. For fixed $t, u(t,x)$ is concave and increasing in x and satisfies the Inada condition at ∞:

$$\lim_{x\to\infty} u'(t,x) = 0$$

However now the agent is diffident about the constant estimates $\hat{\mu}$ and $\hat{\sigma}$, for the drift and volatility matrix respectively and so things change.

We introduce the robust framework by assuming e from now on that **Ω** is the n-dimensional Wiener space of continuous functions, with the natural filtration **F** = $(\mathcal{F}_t)_{0\leq t\leq T}$ . Let K be some fixed compact set of $n\ \times\ n$ symmetric and positive definite matrices, containing $\hat{\Sigma}$ and verifying a uniform ellipticity condition:

$\Sigma\ \in\ K\ \Rightarrow\ \exists a\ >\ 0\, s.t. y'\Sigma y\ \geq\ a^2\|y\|^2 \forall \mathrm{y}\ \in\ \mathbb{R}^n.$

The specification via Σ for the ambiguity in the volatility is in line with empirical practice, as the (instantaneous) covariance matrix $\Sigma$ is the estimated object, and not the volatility σ. The Cholesky factorization offers a one-to-one correspondence between symmetric and positive definite matrices $\Sigma$ and lower triangular matrices σ with positive diagonal elements, so that $\Sigma = \sigma\, \sigma'$ . Therefore, if $(\Sigma_t)_t = (\sigma_t \sigma_t')_t$ , the plausible volatilities are:

$$S = \{\sigma \text{ progr. meas.} \mid \sigma_t(\omega) \text{ is lower triangular, with positive diagonal and } \Sigma_t(\omega) \in K \text{ for all } \omega, t\}.$$

The uncertain drift is also assumed to be progressively measurable. For a given realization of the volatility $\sigma_t(\omega)$, or equivalently of the instantaneous covariance matrix $\Sigma_t(\omega)$ it is allowed to vary in

$$U\epsilon(\Sigma_t(\omega)) = \{u \in \mathbb{R}^n \mid (u - \hat{\mu})'\Sigma_t^{-1}(\omega)(u - \hat{\mu}) \leq \epsilon^2\}$$

that is, in an ellipsoid shaped by $\Sigma_t(\omega)$ ,centered at $\hat{\mu}$ and with constant radius $\epsilon \geq 0$. Let us denote the set of plausible processes by

$$\Upsilon := \{(\mu, \sigma)\, progr.meas. \mid \sigma \in S, \mu_t(\omega) \in U_\epsilon(\Sigma t(\omega))\}$$

Let $S$ be the canonical process $\Omega, \mathbf{F}$ namely $S_T(\omega) = \omega(t)$.

The plausible set P of probabilities is the set of $P_s$ such that S satisfies the SDE:

$$dSt = Diag(St)(\mu_t dt + \sigma_t dW_t^P )$$

in which $W^P$ denotes an n dimensional P-Brownian motion, for some $(\mu, \sigma) \in \Upsilon$

Given the initial wealth $x > 0$, the investment/consumption pair $(\theta\, , c)$ is called (robust) admissible if the measurability, integrability and nonnegativity assumptions hold P-a.s. for all P ∈ $\mathcal{P}$. Namely,

$\mathcal{A}^{rob}(x) := \cap_{P\in\mathcal{P}} \mathcal{A}^P(x)$.

So, the wealth X has P-dynamics given by:

$$dXt = (rXt + \theta'(\hat{\mu} - r\mathbf{1}) - c_t)dt + \theta_t'\, \sigma_t dW_t^P$$

The ambiguity averse investor takes a prudential worst case approach, or alternatively she plays a game against the adverse market. The agent faces now the robust version of the Merton problem:

$$V(0,x) := \sup_{(\theta,c)\in A_{rob}(x)} \inf_{P\in\mathcal{P}} \mathbb{E}^P\left[\int_0^T u(t,c_t)dt + U(T,X_T)\right]$$

Robust verification theorem. Suppose that:

1. there exists a function $v : [0,T] \times \mathbb{R}^+ \to \mathbb{R}$ continuous on $[0,T] \times \mathbb{R}^+$ and $C^{1,2}$ on $[0,T) \times \mathbb{R}^+$ verifying $v(T,\cdot) = U(\cdot)$;

2. for any $(\theta, c) \in A_{rob}(x)$ there exist an optimal $P^{(\theta,c)} \in \mathcal{P}$ of the inner minimization, such that

$$Y_t = Y_t^{(\theta,c)} \equiv v(t, X_t) + \int_0^t u(s, c_s)ds$$

Is a $P^{(\theta,c)}$-supermartingale;

3. there exist some $(\bar{\theta}, \bar{c}) \in A_{rob}(x)$ such that the corresponding $\bar{Y}$ is a $P^{(\bar{\theta},\bar{c})}$ – martingale

Then $(\bar{\theta}, \bar{c}, P^{(\bar{\theta},\bar{c})})$ is an optimizer for the robust Merton problem and $v(0, x) = V(0, x)$.

**HJB-Isaacs equation for the candidate value function:**

Using the Ito's formula under $P \in \mathcal{P}$, the any process Y verifies the SDE:

$$dYt = \left(u(t,c) + v_t + v_x(rx + (\theta'(\hat{\mu} - r\mathbf{1}) - c_t) + \frac{1}{2}v_{xx}\theta'\Sigma\theta\right)dt + v_x\theta_t'\sigma$$

By Ito's Lemma, we derive a drift condition: the sup over the agent's controls of the inf over Nature's controls of Y 's drift must be zero. Thus, a sup-inf non linear PDE arises of HJBI type:

$$\sup_{(\theta,c)} \inf_{(\Sigma,\mu)} \left\{u(t,c) + v_t + v_x(rx + (\theta'(\hat{\mu} - r\mathbf{1}) - c) + \frac{1}{2}v_{xx}\theta'\Sigma\theta\right\} = 0$$

Which is of the HJBI type, where $(\theta, c) \in R^N \times R_+$, and $\Sigma \in K$, $\mu \in U_\epsilon(\Sigma)$ ,

Under the assumptions on *u* and K, the supremum and the infimum in the HJBI equation

$$\sup_{(\theta,c)\in R^n\times R_+} \inf_{\Sigma\in K,\mu\in U_\epsilon(\Sigma)} \left\{u(t,c) + v_t + v_x(rx + (\theta'(\hat{\mu} - r\mathbf{1}) - c) + \frac{1}{2}v_{xx}\theta'\Sigma\theta\right\} = 0$$

Are attained for any $v \in C^{1,2}$ on $(0, T) \times \mathbb{R}^+$ with $v_x > 0, v_{xx} < 0$.

*Proof:* we first minimize first over $\mu \in U_\epsilon(\Sigma)$ for $\Sigma$ fixed.

This amounts to the minimization of the linear function:

$$v_x\theta'\mu$$

over the ellipsoid, and is just an exercise in constrained optimization.

The optimizer is unique when $\theta \neq 0$

$$\mu(\theta) := \hat{\mu} - \epsilon \frac{\Sigma\theta}{\sqrt{\theta'\Sigma\theta}}$$

Substituting it back in the Hamiltonian of the HJBI equation, we get

$$\sup_{(\theta,c)\in R^n\times R_+}\inf_{\Sigma\in K}\left\{u(t,c)+v_t+v_x(rx\ +\left(\theta'(\hat{\mu}-r\mathbf{1})-\ \epsilon\sqrt{\theta'\Sigma\theta}\right)+\frac{1}{2}v_{xx}\theta'\Sigma\theta\right\}=0$$

Which covers also the case θ=0. From this, the minimization over K:

$$\inf_{\Sigma\in K}\left[-\epsilon v_x\ t\sqrt{\theta'\Sigma\theta}+\frac{1}{2}v_{xx}\theta'\Sigma\theta\right]$$

Set $s\ =\ \sqrt{\theta'\Sigma\theta}$. Then the above is the restriction of the concave parabola:

$$y(s)\ =\ -\epsilon v_x\ s\ +\ \frac{1}{2}\ v_{xx}\ s^2$$

to a compact subset of the positive axis. Since the vertex has negative abscissa, the minimum is reached for the maximum s. It follow that the optimizers are those $\Sigma\ \in\ K$ for which $\theta'\Sigma\theta$ is maximal.

Call $M(\theta):=\ \max_K\theta'\ \Sigma\theta$ — continuous function of θ.

The last step is the maximization in the HJBI equation:

$$\operatorname{Sup}_{(\theta,c)\in\mathbb{R}^n\times\mathbb{R}^+}\left[u\ (t,c)+v_t+v_x(rx\ +\theta'(\hat{\mu}-r\mathbf{1})-\epsilon\sqrt{M(\theta)}-c)+\frac{1}{2}M(\theta)v_{xx}\ =\ 0\right]$$

The maximization can be split into the sum of:

1. $\sup_{c\in\mathbb{R}^+}(u(t,c)+v_t-cv_x)$

Concavity of $u\ -\ vx\ >\ 0$ and the Inada on $u$ imply $\lim_{\to+\infty}[u(t,c)\ +\ v_t\ -\ cv_x]=-\infty$.

By continuity, sup is a max.

2. $\sup_{\theta\epsilon\mathbb{R}^n}\ (v_x\left(\theta'(\hat{\mu}-r\mathbf{1})\right)-\epsilon\sqrt{M(\theta)})\ +\ \frac{1}{2}M(\theta)v_{xx})$

$$\mathrm{v}_x\left(\theta'(\hat{\mu}-r\mathbf{1})\right)-\epsilon\sqrt{M(\theta)})\ +\ \frac{1}{2}M(\theta)v_{xx})\ \le$$

$$\le\mathrm{v}_x\left(\theta'(\hat{\mu}-r\mathbf{1})\right)-\epsilon\mathrm{h}\|\theta\|)\ +\ \frac{1}{2}\mathrm{h}^2\|\theta\|^2v_{xx})$$

Coercitivity when $\|\theta\|\ \to\ \infty$. The sup is then attained by some $\bar{\theta}$ since the objective function is continuous.

So let's now analyze what will happen in the case of finite horizon and infinite horizon for non ambiguous σ. The result is based on a max-min Hamilton-Jacobi-Bellman-Isaacs equation. In this section we want to derive a closed form portfolio optimization rule for an investor who is diffident about the drift and has a CRRA utility. The result depends on a max-min Hamilton-Jacobi-Bellman-Isaacs PDE, which broadens the classical Merton problem. As we assume that ambiguity is present only in the drift, as a matter of first importance we have to specify that the lack of uncertainty in the square volatility matrix may be justified by the consideration that mean returns are subject to imprecision to a much higher extent than volatilities. (Biagini, Pinar, 2017).

First, we modify the notation so as to remove the hat over $\hat{\Sigma}$ . The set $\Upsilon$ becomes:

$$\{(\mu,\sigma\ )\mid \mu\, progr.meas.and\ \mu_t(\omega)\ \in\ U_\epsilon(\Sigma)\ for\ all\ \omega\}.$$

We call $\mathcal{P}_\Sigma$the set of plausible probabilities, to underline that Σ is fixed

Thus, we solve and find the explicit solutions to the robust problems for the infinite and finite horizon planning.

## 4.3 The infinite horizon planning for non ambiguous σ:

We assume that the agent has CRRA utility. We fix the positive constants ρ and R, representing the time impatience rate and relative risk aversion, respectively. Follow that the utility is of the form:

$u(t,x) = \ e^{-\rho t}\frac{x^{1-R}}{1-R},$ with $R \neq 1$, or

$u(t,x)\ =\ e^{-\rho t}\ln x$ if $R\ =\ 1.$

In the infinite horizon case, we wish to find the solution of:

$$V_\Sigma(0,x)\ =\ \sup_{(\theta,c)\in A_{rob}(x)}\ \inf_{P\in\mathcal{P}_\Sigma}\ \mathbb{E}^P\,[\,u(s,c_s)\,ds\,]$$

We assume for now that this problem is finite valued, and that both the inner infimum (for a fixed $(\theta\ ,c)\ \in\ A^{rob}(x)$) and the outer supremum are attained. As we seen in the classical case, a guess at the value function takes the form:

$v(t,x)\ =\ (\gamma_\epsilon\ )^{-R}e^{-\rho t}\ \frac{x^{1-R}}{1-R}\ ,R\ \neq 1$ and

$v(t,x)\ =\ e^{-\rho t}\ (\frac{\ln x}{\rho}\ +\ k_\epsilon\ ),R\ =\ 1,$

where $\gamma_\epsilon$ and $k_\epsilon$ are positive constants to be determined. We use $\epsilon$ as subscript to stress t the dependence on the radius of drift ambiguity $\epsilon$ . The candidate $v$ verifies the condition by which: the supremum and the infimum in the HJBI equation are attained for any $v\ \in\ C^{1,2}$ on $(0,T)\times\ \mathbb{R}^+$ with $v_x\ >\ 0, v_{xx} <\ 0,$ so that the optima are attained in the HJBI equation. Substituting the optimal μ(θ) as we stated before with:

$$\mu(\theta) := \hat{\mu} - \epsilon \frac{\Sigma\theta}{\sqrt{\theta'\Sigma\theta}}$$

,

the residual optimization is given by:

$$\sup_{(\theta,c)\in\mathbb{R}^n\times\mathbb{R}^+}\left[u\,(t,c) + v_t + v_x(rx + \theta'(\hat{\mu} - r\mathbf{1}) - \epsilon\sqrt{M(\theta)} - c) + \frac{1}{2}M(\theta)v_{xx} = 0\right]$$

By maximizing over c we obtain:

$\bar{c} = \gamma_\epsilon\, x$ in the power case

$\bar{c} = \rho x$ in the logarithmic case,

And the optimal value is given by:

$$\max_c\{u(t,c) - cv_x\} = e^{-\rho t}\psi_\epsilon(x,R),$$

in which we set:

$\psi_\epsilon\,(x,R) = \frac{R}{1-R}\,(\gamma_\epsilon\, x)^{1-R}, \neq 1$

$\psi_\epsilon\,(x,1) = \ln\rho x - 1\, for\; R = 1.$

Based on what we stated above, the function that we have to maximized is:

$$\max_\theta\left[e^{-\rho t}\,\psi_\epsilon\,(x,R) + v_t + v_x\,(rx + \theta'(\hat{\mu} - r\mathbf{1}) - \epsilon\sqrt{\theta'\Sigma\theta}) + \frac{1}{2}\theta'\Sigma\theta v_{xx}\right]$$

It is concave in θ, and smooth in $R^n \setminus \{0\}$. The first order conditions are thus necessary and sufficient for optimality in $\theta \neq 0$. So, by equating the gradient to zero we obtain:

$$\theta\,(s) = \frac{-svx}{sv_{xx} - vx_\epsilon}\,\Sigma^{-1}(\hat{\mu} - r\mathbf{1})$$

,

where $s := \sqrt{\theta'\,\Sigma\theta}$. We are left with:

$$s^2 = \theta\,(s)'\Sigma\theta(s)$$

Set

$H := \sqrt{(\hat{\mu} - r\mathbf{1})' \Sigma^{-1}((\hat{\mu} - r\mathbf{1})}$, $H_\epsilon := H - \epsilon$

The root of the equation is positive and is given by:

$$\bar{s} = \frac{-v_x H_\epsilon}{v_{xx}}$$

if and only if $H_\epsilon > 0$. If $H_\epsilon \leq 0$, the optimal solution thus is necessarily $\bar{\theta} = 0$. Finally, if $H_\epsilon^+$ denotes the positive part of $H_\epsilon$ we can rewrite the optimal solution in a more compact form both for the power and logarithmic case.

$$\bar{\theta} = x \frac{H_\epsilon^+}{RH} \Sigma^{-1}(\hat{\mu} - r\mathbf{1})$$

The value of the constants $\gamma_\epsilon, k_\epsilon$ is found by substituting these $\bar{c}$ and $\bar{\theta}$ back into:

$$\operatorname*{Sup}_{(\theta,c)\in\mathbb{R}^n\times\mathbb{R}^+} \left[ u(t,c) + v_t + v_x(rx + \theta'(\hat{\mu} - r\mathbf{1}) - \epsilon\sqrt{M(\theta)} - c) + \frac{1}{2} M(\theta) v_{xx} = 0 \right]$$

and we solve the equation. By calculating it we obtain:

$\gamma_\epsilon = \frac{\rho + (R-1)\left(r + \frac{1}{2}\frac{(H_\epsilon^+)^2}{R}\right)}{R}$ and

$k = \frac{1}{\rho^2} \left[ \rho \ln \rho + r - \rho + \frac{(H_\epsilon^+)^2}{2} \right]$

which for $\epsilon = 0$ fall back to the constants $\gamma_0 = \frac{\rho + (R-1)\left[r + \frac{1}{2}\frac{H^2}{R}\right]}{R}$, $k_0 = \frac{1}{\rho^2}\left[\rho \ln \rho + r - \rho + \frac{H^2}{2}\right]$ of the classic cases.

For specification of the well-posedness conditions and the verification that $V_\Sigma(0,x) = v(0.x)$ see [4].

## 4.4 The finite horizon planning for non ambiguous σ

Now the investor has a CRRA power utility both from intertemporal and terminal consumption at time T < ∞:

$$u(t,x) = e^{-\rho t} \frac{x^{1-R}}{1-R} for\ 0 \leq t < T\ and\ u(T,x) = A \frac{x^{1-R}}{1-R}$$

In which A is a fixed positive constant.

We call $\mathcal{P}_\Sigma$ the set of plausible probabilities, to stress that Σ is fixed.

Here, we set the deterministic scaling of the CRRA power utility identical to that of the infinite horizon case. But pay attention that if $e^{-\rho t}$ is replaced by an integrable, positive and deterministic function $h(t)$ what is written below will also holds. We want to find the solution of

$$V_{\Sigma}(0,x) \;=\; \sup_{(\theta,c)\in A_{rob}(x)} \inf_{P\in\mathcal{P}_{\Sigma}} \mathbb{E}^{P}\left[\left[\int_{0}^{T} e^{-\rho\sigma}\frac{{c_s}^{1-R}}{1-R}\right] ds \;+\; A\,\frac{{X_T}^{1-R}}{1-R}\right]$$

Using the scaling properties of the CRRA utility, the guess to the value function takes the form: $v(t,x) = f(t)\frac{x^{1-R}}{1-R}$ for some positive, differentiable function satisfying $f(T) = A$.

Follow that now the HJBI equation is:

$$\max_{(\theta,c)\in\mathbb{R}^{n}\times R^{+}}\left[e^{-\rho t}\frac{c^{1-R}}{1-R} + f'(t)\frac{x^{1-R}}{1-R} + f(t)x^{-R}\left(rx + \theta'(\hat{\mu} - r\mathbf{1}) - \epsilon\sqrt{\theta'\Sigma\theta} - c\right) - \frac{R}{2}\,f(t)x^{-R-1}\theta'\Sigma\theta\right] = \;0.$$

Proceeding as in the previous part, we get

$$\bar{c}(,x) \;=\; x\left(\frac{e^{-\rho t}}{f(t)}\right)^{1/R} \;\bar{\theta} \;=\; x\pi_{\epsilon}.$$

Substituting the above into the HJBI equation results in a first order ODE for $f$:

$$\begin{cases} f'(t) + k_{\epsilon}f(t) + Re^{-\frac{\rho}{R}t}(f(t))^{1-\frac{1}{R}} = 0 \\ f(T) = A \end{cases}$$

With

$$k_{\epsilon} := \;(1 \;-\; R)\;\left(r \;+\; \pi_{\epsilon}'\,(\hat{\mu} \;-\; r\mathbf{1}) \;-\; \epsilon\sqrt{\pi_{\epsilon}'\,\Sigma\pi_{\epsilon}} \;-\; \frac{R}{2}\pi_{\epsilon}'\Sigma\pi_{\epsilon}\right) \;=\; (1 \;-\; R)(r \;+\; \frac{(H_{\epsilon}^{+})}{2R}.$$

With the substitution $f(t) \;=\; g(t)^{R}$ , the ODE can be linearized and easily solved:

$$g(t) \;=\; A^{\frac{1}{R}}exp\left(\frac{k_{\epsilon}}{R}(T \;-\; t)\right) + \;e^{\frac{-k_{\epsilon}}{R}t}\;\int_{t}^{T}\exp\left(\frac{k_{\epsilon}-\rho}{R}s\right)ds$$

.

Obviously for an ambiguity neutral investor with $\epsilon = 0$ we fall back to the finite horizon solution of the Merton problem. (Biagini, Pinar 2017).

# 5. CONCLUSIONS

In this thesis we aim at giving a better understanding of the Merton Problem. Merton mathematically solved the typical question of an investor who wants to maximize the expected utility of the portfolio with his investment strategy by using stochastic control. Merton considered a situation in which the agent has limited choice of investing his wealth in a risky asset and a risk-free one, the agent is risk averse and is objective is to maximize the expected utility of his wealth. Merton found that the optimal allocation strategy is to keep a constant fraction of the wealth in the risky asset. Merton used Ito process to derive Hamilton-Jacobi-Bellman equations. Thus, he provide a solution for finite and infinite horizon Merton Problem. Then in order to address parameter uncertainty issues, we developed a robust max-min version of Merton problem under ambiguity aversions, that is, an agent diffident about mean and return estimates, under ellipsoidal representation of ambiguous parameters. Solutions to the robust problem to both infinite horizon and finite horizon version, under a CRRA utility function, were derived based on a max-min Hamilton-Jacobi-Bellman-Isaacs, a partial differential equation that is central to optimal control theory equation.

# REFERENCES


1. A.Borrelli, *Calcolo Stocastico e Mercati finanziari*, University of Ferrara, 2011/2012

2. I. Karatzas, S. E. Shreve, *Brownian Motion and Stochastic Calculus*, Springer, 1996

3. S. Biagini, *Lecture Notes of Mathematical Methods for Financial Markets*, Department of Economics and Management, Pisa University

4. S. Biagini, M.C.Pinar, *The robust Merton problem of an ambiguity averse investor*, Mathematics and Financial Economics, 2017

5. J.Block, *Stochastic Processes and the Mathematical Finance*, University of Pennsylvania, 2008

6. R.G.Gallagher, *Stochastic Processes: Theory for Application*, M.I.T, 2013

7. K.Sigman, *Notes On Brownian Motion*, Columbia University, 2006

8. V. Krishnan, *Probability and random processes*, Wiley, 2016

9. J.C.Hull, *Options, Futures and Other Derivatives*, University of Toronto, 5th edition, 2002.

10. M. Rubinstein, *Markowitz's "Portfolio Selection": A Fifty-Year Retrospective*, 17 December 2002 https://doi.org/10.1111/1540-6261.00453

11. N. Nkum, A. Lotsi, C. Chapman-Wardy, and K. Doku-Amposah, *Portfolio Optimization Using Minimum Variance Line Approach: (A case study of SSNIT)*,

12. F.J. Fabozzi, P.N. Kolm, A. Pachamanova, D.A., Focardi, S.M *Robust Portfolio: Optimization and Management*, Wiley, New York, 2012

13. G. West, *An Introduction to Modern Portfolio Theory*: Markowitz, CAP-M, APT and Black-Litterman. Financial Modelling Agency 2006

14. S.F. Witt, R.Dobbins. *The Markowitz Contribution to Portfolio Theory*, Managerial Finance, 1979

15. L.C.G. Rogers, *Optimal Investment*, SpringerBriefs in Quantitative Finance, 2013

16. R.Tevzadze, T.Toronjadze, T.Uzunashvili, *Robust utility maximization for a diffusion model with misspecified coefficients* , Finance Stoch , 2013

17. R.C.Merton, *Continuous-Time Finance*, Wiley, 1992

18. M.E. Mangram *A Simplified Perspective of the Markowitz Portfolio Theory*, Global Journal of Business Research, 2013