Continuous-Time Portfolio Optimization

金 含 清

Jin Hanqing

A Thesis Submitted in Partial Fulfillment

of the Requirements for the Degree of Doctor of Philosophy

in

Systems Engineering and Engineering Management

Supervised by:

Professor Xun Yu Zhou

July 2004

To my parents

Acknowledgements

This work would not have been possible without the advice and help of many people. Foremost, I wish to express my deep gratitude to my thesis advisor, Professor Xun Yu Zhou, for his excellent supervision, his constant support and encouragement, many valuable suggestions and fruitful discussions during the last three years. It is him who introduced me the continuous-time mean-variance portfolio selection problem. It is him who encouraged me to continue my research when I was depressed by the difficulty. It is him who told me to solve the problems as clearly and thoroughly as possible. I am also grateful to him for giving me the opportunity to travel and get in touch with distinguished scientists in the fields of mathematical finance and stochastic control theory.

I should also thank the co-authors of one of my paper, Professor S.R. Pliska at University of Illinois at Chicago, and Professor T.R. Bielecki at Northeastern Illinois University. I benefited a lot from their wide eyereach and their magnificent writing. Professor Jia An Yan at the Chinese Academy of Science, the co-author of my other paper, helped me a lot during his visit to Professor Xun Yu Zhou. In those days, I could find the answer for any question in stochastic analysis from his super brain.

I am indebted to two professors who share the same name "Shuzhong"! One is Professor Shuzhong Shi at Peking University, who was my master supervisor at Nankai University. He had taught me a powerful mathematical tool, convex analysis, and led me into the field of mathematical finance. The other is Professor Shuzhong Zhang in our department. His excellent lecture on optimization attracted me deeply. He also gave me some valuable advice when I began to work on my thesis.

Andrew Lim at Berkeley is another professor to whom I want to express my appreciation. His work on mean-variance problem with stochastic parameters in an incomplete market has helped me a lot, and it has been a truly enjoyable experience to discuss with him on ambiguity and robust control when he visited our department at the beginning of this year. Professor Wen Jiang Jiang at the Normal University of Yunnan introduced me the OU-type processes, the quantile method and other tools which were outside of my knowledge. His talent on table tennis and cooking has also benefited me very much.

Although it has been only months since I knew Zuoquan Xu, I have enjoyed very much the brainstorming with him. His talent often improved my efficiency in the research. His simple and honest friendship has made my last months before graduation better.

I feel very happy in SEEM with many friends, especially Chunhong Chiu, Xiaofeng Nie, Rui Li, Lian Yu, Li Zhang, Li Chen, Chunguang Liu, Kui Xu, Min Guo, Xiaolei Yuan, Jun Wang, Peiling Fu, Xiaolin Xu, among many others.

I would like to express my great gratitude and love to my wife, Yan Li, for her long-lasting support and love. She takes care of my life and helps me arrange my working progress.

Last but not least I would like to devote my thanks to my parents, who are the source of my power.

Abstract

This thesis is devoted to the mean-risk portfolio optimization problem in a continuous-time financial market, where we want to minimize the risk of the investment and at the same time ensure that a given expected return level is obtained. Three topics are studied in this thesis.

- The first topic is the mean-variance portfolio selection problem with bankruptcy prohibition in a complete continuous-time market. The problem is completely solved using a decomposition approach. Specifically, when bankruptcy is prohibited, we find that the efficient policy for a mean–variance investor is simply to purchase a European put option that is chosen, according to his or her risk preferences, from a particular class of options. Moreover, we obtain the efficient frontier by a system of parameterized equations.

- The second topic is the mean-variance portfolio selection problem with or without constraints in an incomplete continuous-time market. Four models are discussed: portfolios unconstrained, shorting prohibited, bankruptcy prohibited, and both shorting and bankruptcy prohibited. A duality method is used to solve all the models, and explicit solution are obtained when parameters of the market are all deterministic.

- The third topic is the general mean-risk portfolio selection problem in a complete continuous-time market. In this mean-risk problem, we measure the risk by the expectation of a certain function of the deviation of the terminal payoff from its mean. First of all, the weighted mean-variance problem is solved explicitly. The limit of this weighted mean-variance problem, as the weight on the upside variance goes to zero, is the mean–semivariance problem which is shown to admit *no* optimal solution. This negative result is further generalized to a mean–downside-risk portfolio selection problem where the risk has non-zero value only when the terminal payoff is lower than its mean. Finally, a general model is investigated where the risk function is convex. Sufficient and necessary conditions for the existence of optimal portfolios are given. Moreover, optimal portfolios are obtained when they do exist, and asymptotic optimal portfolios are obtained when optimal portfolios do not exist.

Key Words: continuous-time market, portfolio selection, risk, bankruptcy, shorting, backward stochastic differential equation (BSDE), option, semivariance, downside-risk.

摘　　要

馬科維茨的均值－方差模型得到了廣泛的認可，並于1990年獲得諾貝爾經濟學獎。本文從三個方面討論了連續時間市場上的均值－方差和均值－風險組合優化模型。

第一個方面是完全市場上的均值－方差投資組合優化問題。我們着重討論了帶不破產約束的均值－方差問題。本文利用鞅方法，將這個問題分解為靜態優化問題和復製問題，然后通過求解后面兩個問題，完全求解了帶不破產約束的均值－方差問題。我們發現了在不破產約束下，均值－方差投資組合問題的最優解就是根據投資者的偏好，從某一類歐式賣出期權中，挑選一個出來進行復製。這個問題的組合前沿以參數方程的形式被刻畫出來了。

第二個方面是不完全市場上不帶限製和帶限製的均值－方差投資組合問題。我們討論四個均值－方差投資組合模型，即無約束模型，不賣空模型，不破產模型，及既不賣空又不破產模型。我們用對偶的方法將這些問題轉化為另一類問題。在市場參數確定時，通過這樣的轉化，這四個投資組合模型的最優解被顯式表示出來了。

第三個方面是完全市場上的均值－風險投資組合優化問題。本文討論的風險是用某個函數作用到投資組合未來受益關于其均值的偏差后的期望來度量的。我們從求解均值－加權方差模型開始，先發現了均值－半方差模型不存在非平凡最優解。這個否定的結果可以進一步推廣到一般的均值－下半風險模型。最后我們討論了一般的均值－風險問題。我們找到了這個問題的最優解存在的等價條件。當這個等價條件成立的時候，我們求出投資組合問題的最優解；當這個條件不成立的時候，我們找到了一個漸進最優序列，使得他們的目標值單调地收斂到原問題目標值的下確界。

關鍵詞：　連續時間市場，　投資組合選擇，　風險，　破產，　賣空，　倒向隨機微分方程，　期權，　半方差，　下半風險。

Contents

Notation

In this thesis, we use the following additional notation:

$\mathbf{R}^+$	:	the set of nonnegative real numbers;
$\mathbf{R}^m_+$	:	the set of m-dimension real vectors with nonnegative components;
$\mathbf{R}^-$	:	the set of nonpositive real numbers;
$\mathbf{R}^m_-$	:	the set of m-dimension real vectors with nonpositive components;
$\mathbf{N}$	:	the set of positive integers;
$\mathbf{Q}$	:	the set of rational numbers;
M'	:	the transpose of any vector or matrix M;
$\|M\|$	:	$=\sqrt{\sum_{i,j} m_{ij}^2}$ for any matrix or vector $M = (m_{ij})$;
α^+	:	$= \max\{\alpha, 0\}$ for any real number α;
a^-	:	$= \max\{-a, 0\}$ for any real number a;
$\mathrm{int}(A)$	:	the interior of any set A;
$\mathbf{1}_A$	:	the indicator function of any set A;
$L^2_{\mathcal{F}}(0,T;X)$	:	the set of all stochastic processes $x(\cdot)$ from $[0,T]\times\Omega$ to X, which are $\{\mathcal{F}_t\}_{0\le t\le T}$ progressively measurable and satisfy $E\int_0^T \|x(t)\|^2 dt < +\infty$;
$L^2(\mathcal{F}_T, X)$	:	the set of all X-valued, $\mathcal{F}_T$-measurable random variable x satisfying $E\|x\|^2 < +\infty$.

In this thesis, by convention, all vectors are column vectors.

Chapter 1

Introduction

1.1 Background

Economics is a science for studying how to allocate the scarce resource efficiently. Capital is one of the most important resources, and it is always scarce. Finance is the branch of economics to study how to allocate capital efficiently.

A classical problem in finance is as follows. An investor has some (positive) initial capital, and wants to make the best use of his/her capital. So the investor turns to a financial market, where certain investment instruments such as bonds, stocks, and derivative securities are provided. Now the investor is facing his portfolio optimization problem, that is, how to allocate his capital among these investment instruments.

To help an individual investor, we first need to know more about the financial market: the principal of the investment, and the evolution of the prices of the investment instruments. In other words, we need to model the financial market.

Time and *uncertainty* are two of the most important aspects of a financial market. How to embed them into the evolution of the security prices is essential for modelling the market. The financial market in the real world is so complex. In order to capture the essence of the market, it is necessary to abstract the real market, so that the model can focus on the important elements of the market.

For studying the investment problem, the single-period market is a natural and illuminating model. In such a model, investors can only decide their capital allocations at the beginning of the period, and then evaluate the returns of their investments until the end of the period. The uncertainty of the market is modelled by the randomness of the prices or their returns. In Magill, Quinzil and Quinzii's book [47], they use a matrix to represent the returns as discrete random variables. In most single-period market models, the return of an asset is modelled as a general random variable, typically with normal or lognormal distributions. Markowitz's Nobel prize winning work [41, 42, 43] on mean-variance portfolio selection, which is the most important single-period model, has laid a foundation for modern finance.

Multi-period models are more practical than the single-period ones, In these models, the whole period is divided by a sequence of time spots, and in each time interval between two adjacent time spots, the market is modelled in the same way as in a single-period model. Because of the dynamic evolution of the prices, multi-period models are more than the combination of a sequence of single-period models. In a multi-period model, the uncertainty of the market is built into the evolution of the security prices and the information flow at those time spots. The evolution of the prices is often depicted by the increment of the prices, and the information flow is often given by the historical security prices. Sometimes the Markovian property is required to facilitate the study of these models. Multi-period portfolio selection problem has been studied by Samuelson [57], Hakansson [19], Grauer and Hakansson [18], Pliska [54] and Li and Ng [33].

In continuous-time models, investors can make investment decisions at any time during the investment period. These models are more complicated than the discrete-time ones. They are also different from discrete-time ones, because one cannot simply regard a continuous-time model as the limit of the latter by dividing the investment period into many smaller periods. In 1900, Brownian motion was introduced to study the continuous price processes of securities by Bachelier [2]. But his work had not caught enough eyes until around the 1960's when the stochastic analysis was developed. In 1973 the geometric Brownian motion was used to model the price processes of stocks by Black and Scholes in their seminal work [6]. From then on, using Brownian motion to model the price evolution became standard fare in financial theory. It has been applied in most work on continuous-time finance such as in Merton [45, 46], and Karatzas and Shreve [26]. Today many complete theories have been built on these models, and powerful approaches have been developed to study the financial market.

In the real financial market, transactions of securities involve many factors such as transaction cost, minimal transaction unit, *etc.*. For the convenience of research, some simplification is often made. The following assumptions are generally used, and we will assume that they all hold through out this thesis.

1. Transactions bear no transaction fee and no tax is charged;
2. Investors can buy and sell at any admissible time and admissible amount;
3. Investors can borrow and lend money at the same interest rate;
4. The action of individual investors does not affect the prices of assets in the market.

In some special market models, additional transaction rules may be involved. For example, investors may not short sell securities, or bankruptcy is not allowed. With those additional rules, the market will be more complicated but comes closer to reality.

Another important consideration in investment is to clarify the objective of an investor. Generally, an investor wants to make use of his/her capital as efficiently as possible. But for the efficiency, every individual investor may have his/her own rule, based on which a decision is made. Consequently, we need to model these rules for "efficiency". One of the popular ways to measure the "efficiency" for an individual investor is by *utility* function. Utility is a measure of the satisfaction of an individual on the consumption or income. In the early years of the development of microeconomics, utility was depicted by indifference curves. To the best of my knowledge, the numerical utility, now called utility function, was first studied systematically by von Neumann and Morgenstern [64]. Later it was broadly used in microeconomics, such as in Samuelson's classical work [58].

Utility is an ideal concept in economics, which is often used when consumption is involved. See Merton [45], Karatzas [26], among others. It is well accepted that capital is a non-satiation resource, which means people will always prefer more capital. Therefore it seems that the more return of an investment implies the more "efficiency". But when there is uncertainty in the market, we may not be able to compare the returns of two investments. An alternative way is to compare the expectation of a certain function on the return of the investment, which is called expected utility. In fact, most research involving utility is based on the expected utility when facing uncertainty. The expectation of return itself can also be a criterion in the portfolio optimization problem, which is quite unacceptable because of the existence of uncertainty. As stated by Markowitz in [42], "The expected utility maxim appears reasonable offhand. But so did the expected return maxim ...... Perhaps there is some equally strong reason for decisively rejecting the expected utility maxim as well".

It is arguable whether expected utility is good enough to serve as a criterion in a portfolio optimization problem. As a practical alternative, expected return together with some other criteria has also been broadly studied, for example, the mean-risk multi-objective criterion. It is reasonable that the investors prefer higher expected return and less uncertainty from the investment, and they need to balance the trade-off between the expected return and risk. Although the measurement of risk is rather subjective, there are some widely accepted measures of risk, as well as some axioms for the risk measurement. One of the most famous mean-risk multi-objective frameworks is the mean-variance framework proposed by Markowitz [41], whose work is widely regarded as the beginning of the modern portfolio theory.

There have been a lot of works on continuous-time portfolio selection, most of which use the utility approach. Inspired by the significance and elegance of Markowitz's mean-variance model in single-period market, we study in this thesis the mean-risk portfolio optimization problem in continuous time. We will discuss both complete and incomplete market, with various portfolio constraints. In summary, we will establish the Markowitz theory in continuous time.

1.2 Outline of the thesis

Chapter 2 is devoted to building up the continuous market in which we will carry out our study. In this chapter, we also introduce two important concepts of the financial market, namely arbitrage and completeness. Arbitrage free is a crucial condition for the market to be viable. We try to find the condition as close to the equivalent condition as possible for the arbitrage free condition. Completeness is also studied in this chapter, and a sufficient and necessary condition is introduced under some technical assumption.

We begin the portfolio selection problem in Chapter 3. This chapter is motivated by Zhou and Li [70], in which the Markowitz's model was investigated in a complete continuous-time market. In the conclusion obtained in [70], the optimal portfolio will drive the wealth process to bankruptcy with a strictly positive probability during the investment period. In Chapter 3, we set the bankruptcy prohibition as constraint on the portfolio, and study the corresponding mean-variance portfolio selection problem in a complete market. Martingale method is applied to separate the dynamic portfolio selection problem into a static optimization problem and a portfolio replication problem. Lagrange multipliers method is used to solve the former one. We find that the optimal portfolio under the bankruptcy prohibition is the one that replicates a certain put option. Finally, we obtain the efficient frontier explicitly.

In Chapter 4, we turn to the incomplete market. We go through the similar way as in a complete market, but we encounter two difficulties. One is how to apply the separation approach when the market is incomplete, and hence not all the contingent claims are replicable. The other difficulty is how to solve the static optimization problem separated from the original mean-variance portfolio selection problem. It will be demonstrated that the static optimization problem is far more difficult than that in the complete case. In Chapter 4, we discuss how to deal with these two difficulties in four cases: portfolios unconstrained case, no-shorting case, no bankruptcy case, and neither shorting nor bankruptcy case. When the parameters are all deterministic, we are able to get the explicit optimal solution in each case.

In Chapter 5 we go out of the mean-variance framework to the general mean-risk framework. We start from a weighted mean-variance problem, and then study mean-semivariance problem by regarding it as a limit of the weighted mean-variance problem as the weight on the upside risk converges to zero. We find that other than a trivial case, the mean-semivariance problem in a continuous-time market admits no optimal solution. Furthermore, we generalize this negative result to a mean-downside-risk problem. While in a single-period market, we prove at the end of Chapter 5 that the mean-semivariance problem admits optimal solutions if it admits feasible solutions. Motivated by this surprising conclusion, we then turn to find the equivalent condition for a general mean-risk problem in a continuous-time market to admit optimal solutions. When the existence of optimal solutions is assured, we find at least one of them. When there is no optimal solution, we find a asymptotic optimal solution sequence for the mean-risk problem.

Finally, we conclude this thesis in Chapter 6.

Chapter 2

Continuous-Time Market

2.1 Introduction

皮之不存，毛将焉附？

This is a proverb by XunZi, a Chinese ancient philosopher. It says "where the hair stands upon without skin?".

A financial market is a platform for investment, where there are securities serving as investment instruments. To study an investment problem, we first need to model the market clearly. In order to do that, we need to describe the price processes of securities and specify the trading rules.

The following trading rules are assumed to be true throughout this thesis:

1. **Frictionless Market:** There are no transaction cost or taxes, and all securities are perfectly divisible. Securities can be traded at any time and any admissible amount.

2. **Price-Taker:** Investors' actions will not affect the probability distributions of returns of the available securities.

3. **No Dividend:** No dividend will be paid by securities.

There are different ways to model the price processes of securities and the transaction strategies. In Karatzas and Shreve's book [26], the price processes of risky assets are modelled by *stochastic differential equations (SDE)* driven by an n-dimensional Brownian motion; the price process of riskless asset (bond) is modelled by an ordinary differential equation (ODE). And the transaction strategies modelled by the so-called "tame portfolio". With the "tame portfolio", a wealth process can be bounded from below by some constant multiplying the price of bond. In their model, arbitrage-free and completeness of the financial market can be depicted in almost equivalent ways. In Yan and Xia's recent work [65], a more general setup is proposed, where an d-dimensional semimartingale is used to model the price processes of the securities. They define the transaction strategies as such where the corresponding wealth processes are bounded from below by the inner product of a constant vector and the price vector of the securities. They apply the integration of semimartingale to study the market, and build up a set of theorems even though the market is very general.

In this thesis, we will work on the following L^2-system.

2.2 L^2- system

$T > 0$ is a fixed terminal time and $(\Omega, \mathcal{F}, P, \{\mathcal{F}_t\}_{t\geq 0})$ is a filtered complete probability space on which a standard $\mathcal{F}_t$-progressively measurable, m-dimensional Brownian motion $W(t) \equiv (W^1(t), \cdots, W^m(t))'$ with $W(0) = 0$ is defined. It is assumed that the filtration $\{\mathcal{F}_t : 0 \leq t \leq T\}$ is generated by the Brownian motion and augmented by all the P-null sets. We denote by $L^2_{\mathcal{F}}(0, T; \mathbf{R}^d)$ the set of all $\mathbf{R}^d$-valued, $\mathcal{F}_t$-progressively measurable stochastic processes $f(\cdot) = \{f(t) : 0 \leq t \leq T\}$ such that $E\int_0^T |f(t)|^2 dt < +\infty$, and by $L^2(\mathcal{F}_T, \mathbf{R}^d)$ the set of all $\mathbf{R}^d$-valued, $\mathcal{F}_T$-measurable random variables η such that $E|\eta|^2 < +\infty$. Throughout this thesis, a.s. (the abbreviation of "almost surely") signifies that the corresponding statement holds true with probability 1 (with respect to P).

Suppose in the market $n + 1$ assets (or securities) are traded continuously. One of the assets is the *bank account* whose price process $S_0(t)$ is subject to the following (stochastic) ordinary differential equation:

$$\begin{cases} dS_0(t) = r(t)S_0(t)dt, \\ S_0(0) = s_0 > 0, \end{cases} \tag{2.1}$$

where the *interest rate* $r(t)$ is a uniformly bounded, $\mathcal{F}_t$-adapted, scalar-valued stochastic process. Normally in practice, the interest rate $r(t) \geq 0$, yet this assumption is not necessary in the model. The other n assets are *stocks* whose price processes $S_i(t)$, $i = 1, \cdots, n$, satisfy the following stochastic differential equation:

$$\begin{cases} dS_i(t) = S_i(t)\big[b_i(t)dt + \sum_{j=1}^m \sigma_{ij}(t)dW^j(t)\big], \\ S_i(0) = s_i > 0, \end{cases} \tag{2.2}$$

where $b_i(t)$ and $\sigma_{ij}(t)$, the *appreciation* and *diffusion* (or *volatility*) rates, respectively, are scalar-valued, $\mathcal{F}_t$-adapted, uniformly bounded stochastic processes.

Define the *volatility matrix* $\sigma(t) := (\sigma_{ij}(t))_{n\times m}$.

Consider an investor who invests his capital in the market in the following way: at time $t \geq 0$, the amount invested in the bank account is $\pi^0(t)$ and the amount invested in security i is $\pi^i(t)$. We call $(\pi^0(t), \pi^1(t), \cdots, \pi^n(t))'$ a *trading strategy*. Denote his total wealth at time $t \geq 0$ by $x(t)$. Then obviously $x(t) = \sum_{i=0}^n \pi^i(t)$.

We call a trading strategy $(\pi^0(t), \pi^1(t), \cdots, \pi^n(t))'$ to be *self-financing* if and only if

$$dx(t) := d\sum_{i=0}^n \pi^i(t) = \sum_{i=0}^n \frac{\pi^i(t)}{S_i(t)} dS_i(t).$$

Self-financing condition can be understood as that there is no capital injection into nor withdrawal from the investment.

Assume that trading of shares takes place continuously in a self-financing manner. Then $x(\cdot)$ satisfies (see, e.g., Karatzas and Shreve [26] and Elliott and Kopp [14])

$$\begin{cases} dx(t) &= \Big\{r(t)x(t) + \sum_{i=1}^n \big[b_i(t) - r(t)\big]\pi^i(t)\Big\}dt \\ &\quad + \sum_{j=1}^m \sum_{i=1}^n \sigma_{ij}(t)\pi^i(t)dW^j(t), \\ x(0) &= x_0 \geq 0. \end{cases} \tag{2.3}$$

Notice that for a self-financing strategy, $\pi^0(t)$ is determined by $(\pi^1(t), \cdots, \pi^n(t))'$ via (2.3). So we can just use $\pi(t) = (\pi^1(t), \cdots, \pi^n(t))'$ to represent a self-financing strategy.

Set

$$B(t) := (b_1(t) - r(t), \cdots, b_n(t) - r(t))'. \tag{2.4}$$

With this notation, equation (2.3) becomes

$$\begin{cases} dx(t) = [r(t)x(t) + \pi(t)'B(t)]dt + \pi(t)'\sigma(t)dW(t), \\ x(0) = x_0. \end{cases} \tag{2.5}$$

Now we need to define an "allowable" trading strategy.

Definition 2.2.1 A self-financing trading strategy $\pi(\cdot)$ is said to be *admissible* if $\sigma(\cdot)'\pi(\cdot) \in L^2_{\mathcal{F}}(0,T;\mathbf{R}^m)$. The set of all the admissible self-financing trading strategies is denoted by Π.

We also call an admissible self-financing trading strategy as a *portfolio*.

Observe that by the standard SDE theory, a unique strong solution exists for the wealth equation (2.5) for any portfolio $\pi(\cdot)$.

We call the system described above the L^2-system, because all the admissible portfolios $\pi(\cdot)$ pre-multiplied by the volatility matrix $\sigma(\cdot)'$ are in $L^2_{\mathcal{F}}(0,T,\mathbf{R}^m)$, and as well as their corresponding wealth processes.

In the following two sections, we will study two important concepts of the financial market: *arbitrage* and *completeness*.

2.3 Arbitrage

Arbitrage-free is a fundamental concept in an economy market, especially in a financial market.

Definition 2.3.1 A portfolio $\pi(\cdot)$ is called an *arbitrage opportunity* if there exists an initial $x_0 \leq 0$ and a time $t \in [0,T]$, such that the corresponding wealth process $x(\cdot)$ satisfies $P(x(t) \geq 0) = 1$, and $P(x(t) > 0) > 0$.

Definition 2.3.2 A market is called *arbitrage-free* if there is no arbitrage opportunity.

Remark 2.3.1 It is easy to see that a market is not arbitrage-free if and only if there is a portfolio $\pi(\cdot)$ and an initial wealth $x(0) = x_0 \leq 0$ such that the corresponding wealth process satisfies $x(T) \geq 0,$ a.s., and $P\{x(T) > 0\} > 0$.

Absence of arbitrage opportunity is a necessary condition for a financial market to be viable. If there are arbitrage opportunities in the market, then investors can make money from the market as much as they want without paying for anything.

Next we give a sufficient condition and a necessary condition for a market being arbitrage-free.

Theorem 2.3.1 *If there is no $\mathcal{F}_t$-progressively measurable process $\theta(\cdot)$ such that $B(t) = \sigma(t)\theta(t)$,* a.s., *a.e.$t \in [0, T]$, then the market is not arbitrage-free.*

Proof: By Theorem A.1 and Lemma B.2, we can prove that there exists a progressively measurable process $\theta_0(\cdot)$ such that $\theta_0(t) \in \text{argmin}_{x\in\mathbf{R}^m}|\sigma(t)x - B(\cdot)|^2$. Define $\pi_0(\cdot) = \sigma(\cdot)\theta_0(\cdot) - B(\cdot)$, then $\sigma(t)'\pi_0(t) = 0$, $\pi_0(t)'B(t) = |\pi_0(t)|^2$, which imply that $\pi_0(\cdot)$ is a portfolio.

Define $x(\cdot)$ as the wealth process of $\pi_0(\cdot)$ with the initial wealth $x(0) = 0$, then

$$\begin{aligned} dx(t) &= r(t)x(t)dt + \pi_0(t)'[B(t)dt + \sigma(t)dW(t)] \\ &= r(t)x(t)dt + |\pi_0(t)|^2 dt, \\ x(t) &= \int_0^T e^{\int_t^T r(s)ds}|\pi_0(t)|^2 dt \\ &\geq 0. \end{aligned}$$

Define $\underline{r}$ to be a lower bound of the unique bounded process $r(\cdot)$. If there does not exist a $\theta(\cdot)$ satisfies $\sigma(t)\theta(t) = B(t)$ a.s., *a.e.$t \in [0, T]$*, then $P\{\int_0^T |\pi_0(t)|^2 dt > 0\} > 0$. However,

$$\begin{aligned} x(T) &= \int_0^T e^{\int_t^T r(s)ds}|\pi_0(t)|^2 dt \\ &\geq e^{\min\{\underline{r}T,0\}} \int_0^T |\pi_0(t)|^2 ds. \end{aligned}$$

Hence,

$$\begin{aligned} P\{x(T) > 0\} &\geq P\{\int_0^T |\pi_0(t)|^2 dt > 0\} \\ &> 0. \end{aligned}$$

This means that the market is not arbitrage-free. □

Theorem 2.3.2 *If there exists a uniformly bounded $\mathcal{F}_t$-adapted process $\theta(\cdot)$, such that $B(t) = \sigma(t)\theta(t)$* a.s. *a.e.$t \in [0, T]$, then the market is arbitrage-free.*

Proof: By Remark 2.3.1, we need only to prove that for any wealth process $x(\cdot)$ under a portfolio $\pi(\cdot)$, if $x(T) \geq 0$, a.s. and $P\{x(T) > 0\} > 0$ then $x(0) \geq 0$.

Let us fix an admissible portfolio $\pi(\cdot)$ and let $x(\cdot)$ be the unique wealth process that solves (2.5) and $x(T) \geq 0$, a.s. and $P\{x(T) > 0\} > 0$. Note that $\xi := x(T)$ is a positive square-integrable $\mathcal{F}_T$-measurable random variable; hence $(x(\cdot), z(\cdot)) := (x(\cdot), \sigma(\cdot)'\pi(\cdot))$ satisfies the following *backward stochastic differential equation (BSDE)*:

$$\begin{cases} dx(t) = [r(t)x(t) + z(t)'\theta(t)]dt + z(t)'dW(t), \\ x(T) = \xi. \end{cases} \tag{2.6}$$

Applying Proposition 2.2 (Page 22) in El Karoui, Peng and Quenez [13], we obtain the following representation

$$x(t) = \rho(t)^{-1}E(\rho(T)x(T)|\mathcal{F}_t), \quad \forall t \in [0,T], \text{ a.s.,} \tag{2.7}$$

where $\rho(\cdot)$ satisfies

$$\begin{cases} d\rho(t) = \rho(t)[-r(t)dt - \theta(t)'dW(t)], \\ \rho(0) = 1, \end{cases} \tag{2.8}$$

or, equivalently,

$$\rho(t) = \exp\left\{-\int_0^t [r(s) + \frac{1}{2}|\theta(s)|^2]ds - \int_0^t \theta(s)'dW(s)\right\}. \tag{2.9}$$

It follows from (2.7) that $\forall t \in [0,T]$, $x(t) > 0$, a.s.. In particular, $x_0 = x(0) > 0$, which implies that there is no arbitrage opportunity in the market. $\square$

Observe that the above process $\rho(\cdot)$ in (2.9) is nothing else but what financial economists call the *deflator* process. Since for our market, under the condition of Theorem 2.3.2, there exists an equivalent martingale measure Q, satisfying

$$\frac{dQ}{dP}\Big|_{\mathcal{F}_t} = \eta(t), \quad \text{a.s.,}$$

where $\eta(t) := \frac{S_0(t)}{s_0}\rho(t)$. Thus representation (2.7) can be rewritten as the *risk-neutral* valuation formula

$$x(t) = S_0(t)E_Q[S_0(T)^{-1}x(T)|\mathcal{F}_t], \quad \forall t \in [0,T], \text{ a.s.,}$$

where we denote by E_Q the expectation with respect to the probability Q.

It should be noted that in Theorem 2.3.2, the requirement that $\theta(\cdot)$ is uniformly bounded is strong, and it is difficult to be weakened. However, for a market with deterministic parameters, we can weaken it to $\int_0^T |\theta(t)|^2 dt < +\infty$.

Theorem 2.3.3 *If* $r(\cdot), b(\cdot), \sigma(\cdot)$ *are all deterministic, then the market is arbitrage-free if there exists a (deterministic) process* $\theta(\cdot)$ *such that* $B(t) = \sigma(t)\theta(t)$ a.s. $a.e. t \in [0, T]$ *with* $\int_0^T |\theta(t)|^2 dt < +\infty$.

Proof: Without loss of generality, suppose $r(\cdot) = 0$. (Else we can study $\bar{x}(\cdot) = e^{\int_0^\cdot r(s)ds} x(\cdot)$ instead of $x(\cdot)$). Thanks to $\int_0^T |\theta(s)|^2 ds < +\infty$, we have $e^{\int_0^T |\theta(s)|^2 ds} < +\infty$. By Novinkov's condition, we know $\rho(t) := e^{-\int_0^t |\theta(s)|^2 ds/2 - \int_0^t \theta(s)' dW(s)}$, $t \in [0, T]$, is a martingale.

Take $\theta_k(t) = \text{cap}(\theta(t), k)$, where for any $x \in \mathbf{R}^n$, $\text{cap}(x, k)$ is defined as the vector $y = (y_1, \cdots, y_n)'$ with $y_i = \text{sgn}(x_i) \min\{|x_i|, k\}$. Then $\theta_k(\cdot)$ is bounded by k, and $\theta_k(\cdot) \xrightarrow{L^2} \theta(\cdot)$.

For any x_0 and $\pi(\cdot)$, define

$$\left\{ \begin{array}{rcl} dx(t) & = & \pi(t)'[B(t)dt + \sigma(t)dW(t)] \\ & = & [\sigma(t)'\pi(t)]'[\theta(t)dt + dW(t)], \\ x(t) & = & x_0, \end{array} \right.$$

$$\left\{ \begin{array}{rcl} dx_k(t) & = & [\sigma(t)'\pi(t)]'[\theta_k(t)dt + dW(t)], \\ x_k(0) & = & x_0. \end{array} \right.$$

Denote $D_k(t) = x_k(t) - x(t)$. It is easy to see

$$\left\{ \begin{array}{rcl} dD_k(t) & = & [\sigma(t)'\pi(t)]'[\theta_k(t) - \theta(t)]dt, \\ D_k(0) & = & 0. \end{array} \right.$$

Therefore

$$\begin{array}{rcl} 0 & \leq & ED_k(T)^2 \\ & = & E\left\{ \int_0^T [\sigma(t)'\pi(t)]'[\theta_k(t) - \theta(t)]dt \right\}^2 \\ & \leq & E\left\{ \int_0^T |\sigma(t)'\pi(t)|^2 dt \int_0^T |\theta_k(t) - \theta(t)|^2 dt \right\} \\ & = & \int_0^T |\theta_k(t) - \theta(t)|^2 dt \, E \int_0^T |\sigma(t)'\pi(t)|^2 dt. \end{array}$$

From the facts that $\theta_k(\cdot) \xrightarrow{L^2} \theta(\cdot)$ and $E\int_0^T |\sigma(t)'\pi(t)|^2 dt < +\infty$, it follows that the last one converges to 0. So $E(x_k(T) - x(t))^2 \to 0$.

On the other hand, denote $\rho_k(t) = e^{-\int_0^T |\theta_k(t)|^2/2 - \int_0^T \theta_k(t)' dW(t)}$. A gain by Proposition 2.2 (Page 22) in El Karoui, Peng and Quenez [13], we obtain that $E[x_k(T)\rho_k(T)] = x_k(0) = x_0$. Furthermore,

$$\begin{aligned} E(\rho_k(T) - \rho(T))^2 &= E\rho_k(T)^2 + E\rho(T)^2 - 2E[\rho_k(T)\rho(T)] \\ &= e^{\int_0^T |\theta_k(t)|^2 dt} + e^{\int_0^T |\theta(t)|^2 dt} - 2e^{\int_0^T \theta_k(t)'\theta(t) dt} \\ &\to 0. \end{aligned}$$

Therefore $E[x_k(T)\rho_k(T)] \to E[x(T)\rho(T)]$, implying $E[x(T)\rho(T)] = x_0$. Therefore the market is arbitrage-free. $\square$

2.4 Completeness

The completeness of the market to be introduced here is different from the one in general finance theory. Let us start with the definition.

Definition 2.4.1 A European contingent $\xi \in L^2(\mathcal{F}_T, \mathbf{R})$ is said to be *replicable* if there exist an initial wealth x_0 and a portfolio $\pi(\cdot)$ such that the wealth of the portfolio $x(\cdot)$ satisfies $x(T) = \xi$. A market is called *complete* if any contingent claim $\xi \in L^2(\mathcal{F}_T, \mathbf{R})$ is replicable.

Completeness is another important notion for a market. A market being complete means that the investors can obtain any (square integrable) contingent return at time T. The completeness of a market can help us handle the market easier; but in most cases, the market is incomplete.

Similar to the arbitrage-free condition, we are not yet able to obtain an equivalent technical condition for the completeness in general. But with some additional assumption, we can find a sufficient and necessary condition for the market to be complete.

Proposition 2.4.1 *Suppose there exists a uniformly bounded $\theta(\cdot)$ such that $\sigma(t)\theta(t) = B(t)$,* a.s., *$a.e. t \in [0, T]$. Then the market is complete if and only if* $\text{rank}(\sigma(t)) = m$, a.s., *$a.e. t \in [0, T]$.*

Proof: Without loss of generality, suppose $r(t) \equiv 0$. Denote $\bar{\pi}(t) = \sigma(t)'\pi(t)$. Then we can rewrite the wealth equation (2.5) as

$$dx(t) = \bar{\pi}(t)'[\theta(t)dt + dW(t)].$$

If $\text{rank}(\sigma(t)) = m$, a.s., $a.e.t \in [0, T]$, then for any contingent claim $\xi \in L^2(\mathcal{F}_T, \mathbf{R})$, the BSDE

$$\begin{cases} dx(t) = \bar{\pi}(t)'[\theta(t)dt + dW(t)], \\ x(T) = \xi. \end{cases}$$

admits a unique solution pair $(x(\cdot), \bar{\pi}(\cdot)) \in L^2_{\mathcal{F}}(0, T, \mathbf{R}) \times L^2_{\mathcal{F}}(0, T, \mathbf{R}^m)$. Because $\text{rank}(\sigma(t)) = m$, a.s. $a.e.t \in [0, T]$, we know there exists $\pi(t) \in \mathbf{R}^n$ such that $\sigma'(t)\pi(t) = \bar{\pi}(t)$, which implies ξ is replicable.

On the other hand, for any $\bar{\pi}(\cdot) \in L^2_{\mathcal{F}}(0, T, \mathbf{R}^m)$, denote

$$x(t) = \int_0^t \bar{\pi}(s)'\theta(s)ds + \int_0^t \bar{\pi}(s)'dW(s).$$

If the market is complete, then $x(T)$ is replicable, that is, there exists a $\pi(\cdot) \in L^2_{\mathcal{F}}(0, T, \mathbf{R}^n)$ such that $\sigma(t)'\pi(t) \in L^2_{\mathcal{F}}(0, T, \mathbf{R}^m)$ and

$$x(t) = \int_0^t \pi(s)'\sigma(s)\theta(s)ds + \int_0^t \pi(s)'\sigma(s)dW(s).$$

By the uniqueness of the solution for BSDE

$$\begin{cases} dy(t) = \bar{\pi}(t)'[\theta(t)dt + dW(t)], \\ y(T) = \xi. \end{cases}$$

we know $\bar{\pi}(t) = \sigma(t)'\pi(t)$, which implies that $\text{rank}(\sigma(t)) = m$. □

2.5 Remarks on L^2-system

L^2-system is a convenient system for studying the portfolio optimization problem, where the martingale method and backward stochastic differential equation theory can be used confidently. A major advantage in using this system, as will be shown later in this thesis, is that the set of all the replicable European contingent claims can be depicted explicitly, which provides a strong base for the dual method.

As we see in Sections 2.3 and 2.4, many conclusions rely on the assumption that $r(\cdot), B(\cdot), \sigma(\cdot)$ be uniformly bounded. Furthermore, the assumption that there exists a uniformly bounded process $\theta(\cdot)$ such that $\sigma(t)\theta(t) = B(t)$ a.s., $a.e.t \in [0, T]$ plays an essential role in the study of the market. These two assumptions are set for martingale method to work smoothly. In fact, under these assumptions, one can solve some complicate problems in the market. These assumptions can be weakened by some lower bounded condition, for instance the "tameness" in Karatzas and Shreve [26] and the "allowance" in Xia and Yan [65]. But the corresponding results are also weakened.

Chapter 3

Mean-Variance Criteria in a Complete Market

3.1 Introduction

Mean–variance portfolio selection is concerned with the allocation of wealth among a variety of securities so as to achieve the optimal trade-off between the expected return of the investment and its risk over a fixed planning horizon. The model was first proposed and solved more than fifty years ago in the single-period setting by Markowitz in his Nobel-Prize winning work Markowitz [41], [42]. With the risk of a portfolio measured by the variance of its return, Markowitz showed how to formulate the problem of minimizing a portfolio's variance subject to the constraint that its expected return equals a prescribed level as a quadratic program. Such an optimal portfolio is said to be *variance minimizing*, and if it also achieves the maximum expected return among all portfolios having the same variance of return, then it is said to be *efficient*. The set of all points in the two-dimensional plane of variance (or standard deviation) and expected return that are produced by efficient portfolios is called the *efficient frontier*. Hence investors should focus on the efficient frontier, with different investors selecting different efficient portfolios, depending upon their risk preferences.

Not only does this model and its single-period variations (e.g., there might be constraints on the investments in individual assets) witness widespread use in the financial industry, but also the basic concepts underlying this model have become the cornerstone of classical financial theory. For example, in Markowitz's world (i.e., the world where all the investors act in accordance with the single-period, mean–variance theory), one of the important consequences is the so-called mutual fund theorem, which asserts that two mutual funds, both of which are efficient portfolios, can be established so that all investors will be content to divide their assets between these two funds. Moreover, if a risk-free asset (such as a bank account) is available, then it can serve as one of the two mutual funds. A logical consequence of this is that the other mutual fund, which itself is efficient, must correspond to the "market." This, in turn, leads to the elegant *capital asset pricing model* (CAPM), see Sharpe [61], Lintner [37], Mossin [48].

Meanwhile, in subsequent years there has been considerable development of multi-period and, pioneered by the famous work Merton [45], continuous-time models for portfolio management. In these work, however, the approach is considerably different, as expected utility criteria are employed. For example, for the problem of maximizing the expected utility of the investor's wealth at a fixed planning horizon, Merton [45] used dynamic programming and partial differential equation theory to derive and analyze the relevant Hamilton–Jacobi–Bellman (HJB) equation. The recent books by Karatzas and Shreve [26] and Korn [29] summarize much of this continuous time, portfolio management theory.

Multi-period, discrete-time mean–variance portfolio selection has been studied by Samuelson [57], Hakansson [19], Grauer and Hakansson [18], and Pliska [54]. But somewhat surprisingly, the *exact, faithful* continuous-time versions of the mean–variance problem have not been developed until very recently. This is surprising because the mean–variance portfolio problem is known to be very similar to the problem of maximizing the expected *quadratic* utility of terminal wealth. Solving the expected quadratic utility problem can produce a point on the mean–variance efficient frontier, although *a priori* it is often unclear what the portfolio's expected return will turn out to be. So while it is straightforward to formulate a continuous-time version of the mean–variance problem as a dynamic programming problem, researchers have been slow to produce significant results.

A more modern approach to continuous-time portfolio management, first introduced by Pliska [52],[53], avoids dynamic programming by using the risk neutral (martingale) probability measure; but this has not been much helpful either. This risk neutral computational approach decomposes the problem into two sub-problems, where first one uses convex optimization theory to find the random variable representing the optimal terminal wealth, then solves the sub-problem of finding the trading strategy that replicates the terminal wealth. The solution for the mean–variance problem of the first sub-problem is known for the unconstrained case, [1] but apparently nobody has successfully solved for continuous time applications the second sub-problem, which is essentially a martingale representation problem.

A breakthrough of sorts was provided in a recent paper by Li and Ng [33], who studied the discrete-time, multi-period, mean–variance problem using the framework of multi-objective optimization, where the variance of the terminal wealth and its expectation are viewed as competing objectives. They are combined in a particular way to give a single-objective "cost" for the problem. An important feature of this paper is an embedding technique, introduced because dynamic programming could not be directly used to deal with their particular cost functional. Their embedding technique was used to transform their problem to one where dynamic programming was used to obtain explicit optimal solutions.

Zhou and Li [70] used the embedding technique and linear–quadratic (LQ) optimal control theory to solve the continuous-time, mean–variance problem with assets having deterministic diffusion coefficients. In their LQ formulation, the dollar amounts, rather than the proportions of wealth, in individual assets are used to define the trading strategy. This leads to a dynamic system that is linear in both the state (i.e., the level of wealth) and the control (i.e., the trading strategy) variables. Together with the quadratic form of the objective function, this formulation falls naturally into the realm of stochastic LQ control. Moreover, since there is no running cost in the objective function, the resulting

[1] See, for example, Pliska [54]; the treatment there was for the single-period situation, but the basic result easily generalizes to very similar results for the multi-period and continuous-time situations.

problem is inherently an *indefinite stochastic LQ control problem*, the theory of which has been developed only very recently (see, e.g., [67, Chapter 6]).

Exploiting the stochastic LQ control theory, Zhou and his colleagues have considerably extended the initial continuous-time, mean–variance results obtained by Zhou and Li [70]. Lim and Zhou [36] allowed for stocks which are modeled by processes having random drift and diffusion coefficients, Zhou and Yin [71] featured assets in a regime switching market, and Li, Zhou, and Lim [34] introduced a constraint on short selling. Kohlmann and Zhou [27] went in a slightly different direction, studying the problem of mean–variance hedging of a given contingent claim. In all these papers, explicit forms of efficient/optimal portfolios and efficient frontiers were presented. While many results in the continuous-time Markowitz world are analogous to their single-period counterparts, there are some results that are strikingly different. Most of these results are summarized by Zhou [69], who also provided a number of examples that illustrate the similarities as well as differences between the continuous-time and single-period settings.

In view of all this recent work on the continuous-time, mean–variance problem, what is left to be done? The answer is that it is desirable to address a significant shortcoming of the preceding models, for their resulting optimal trading strategies can cause bankruptcy for the investor. Moreover, these models assume a bankrupt investor can keep on trading, borrowing money even though his or her wealth is negative. In most of the portfolio optimization literature the trading strategies are expressed as the proportions of wealth in the individual assets, so with technical assumptions (such as finiteness of the integration of a portfolio) about these strategies the portfolio's monetary value will automatically be *strictly* positive. But with strategies described by the money invested in individual assets, as dictated by the stochastic LQ control theory approach, a *larger* set of trading strategies is available, including ones which allow the portfolio's value to reach zero or to become and remain strictly negative (e.g., borrow from the bank, buy stock on margin, and watch the stock's price go into the tank). The ability to continue trading even though the value of an investor's portfolio is strictly negative is highly unrealistic. This brings us to the subject of this chapter: the study of the continuous-time, mean–variance problem with the additional restriction that bankruptcy is prohibited[2].

[2]Here the bankruptcy is defined as the wealth being *strictly* negative. A zero wealth is *not* regarded

In this chapter we use an extension of the risk neutral approach rather than making heavy use of stochastic LQ control theory. However, we retain the specification of trading strategies in terms of the monetary amounts invested in individual assets, and we add the explicit constraint that feasible strategies must be such that the corresponding monetary value of the portfolio is nonnegative (rather than strictly positive) at every point in time with probability one. The resulting continuous time, mean–variance portfolio selection problem is straightforward to formulate, as will be seen in the following section. Our model of the securities market is complete, although we allow the asset drift and diffusion coefficients, as well as the interest rate for the bank account, to be random. Once again, we emphasize that the set of trading strategies we consider is larger than that of the proportional strategies, and we will show that the efficient strategies we obtain are in general *not* obtainable by the proportional ones. In Section 3.2 we also demonstrate that the original nonnegativity constraint can be replaced by the constraint which simply requires the terminal monetary value of the portfolio to be nonnegative. This leads to the first sub-problem in the risk neutral computational approach: find the nonnegative random variable having minimum variance and satisfying two constraints, one calling for the expectation of this random variable under the original probability measure to equal a specified value, and the other calling for the expectation of the *discounted* value of this random variable under the risk neutral measure to equal the initial wealth.

In Section 3.3 we study the feasibility of our problem, an issue that has never been addressed by other authors to the best of our knowledge. There we provide two nonnegative numbers with the property that the variance minimizing problem has a unique, optimal solution if and only if the ratio of the initial wealth to the desired expected wealth falls between these two numbers. In Section 3.4 we solve the first sub-problem by introducing two Lagrange multipliers that enable the problem to be transformed to one where the only constraint is that the random variable, i.e., the terminal wealth, must be nonnegative. This leads to an explicit expression for the optimal random variable, an expression that is in terms of the two Lagrange multipliers which must, in turn, satisfy a system of two equations. In Section 3.5 we show that this system has a unique solution, and we

as in bankruptcy. In fact, as will be seen in the sequel the wealth process associated with an efficient portfolio may indeed "touch" zero with a positive probability.

establish simple conditions for determining what the signs of the Lagrange multipliers. A consequence here is the observation that the optimal terminal wealth can be interpreted as the payoff of either, depending on the signs of the Lagrange multipliers, a European put or a call on a fictitious security.

In Section 3.6 we turn to the second sub-problem, showing that the optimal trading strategy of the variance minimizing problem can be expressed in terms of the solution of a backward stochastic differential equation. We also provide an explicit characterization of the mean–variance efficient frontier, which is a proper portion of the variance minimizing frontier. Unlike the situation where bankruptcy is allowed, the expected wealth on the efficient frontier is not necessarily a linear function of the standard deviation of the wealth. In Section 3.7 we consider the special case where the interest rate and the risk premium are deterministic functions of time (if not constants). Here we provide explicit expressions for the Lagrange multipliers, the optimal trading strategies, and the efficient frontier. We conclude in Section 3.8 with some remarks.

Somewhat related to our work are the continuous-time studies of mean–variance hedging by Duffie and Richardson [12], and Schweizer [59]. More pertinent is the study of continuous-time, mean–variance portfolio selection in Richardson [55], a study where the portfolio's monetary value was allowed to become strictly negative. Also in the working paper of Zhao and Ziemba [68], a mean–variance portfolio selection problem with deterministic market coefficients and with bankruptcy allowed is solved using a martingale approach. Closely connected to our research is the work by Korn and Trautmann [30] and Korn [29]. They considered the continuous-time mean–variance portfolio selection problem with nonnegativity constraints on the terminal wealth for the case of the Black–Scholes market where there is a single risky asset that is modelled as simple geometric Brownian motion and where the bank account has a constant interest rate. They provided expressions for the optimal terminal wealth as well as the optimal trading strategy using a duality method. Their first sub-problem fixes a single Lagrange multiplier and then solves an unconstrained convex optimization problem for the optimal proportional strategy. Their second sub-problem is to find the "correct" value of their Lagrange multiplier. Actually, they do not have an explicit constraint for nonnegative wealth, but by using strategies that are in terms of proportions of wealth, a strictly positive wealth is automatically achieved. In our paper we include strategies that allow the wealth to become zero at intermediate dates, so apparently our set of feasible strategies is larger. Our results are considerably more general, for we allow stochastic interest rates, an arbitrary number of assets, and asset drift and diffusion coefficients that are random. And we provide characterizations of efficient frontiers, necessary and sufficient conditions for existence of solutions, and several other kinds of results that Korn and Trautmann [30] did not address at all.

3.2 Problem formulation

We adopt the market model given in Chapter 2. In this chapter, we suppose the number of risk securities, n, is equal to m, the dimension of the Brownian motion. In addition, we make the following basic assumption throughout this chapter

Assumption 3.2.1

$$\sigma(t)\sigma(t)' \geq \delta I_m, \quad \forall t \in [0,T], \ \text{ a.s.,} \tag{3.1}$$

for some $\delta > 0$, where I_m is the $m \times m$ identity matrix.

By Theorem 2.3.1 and Theorem 2.4.1, the market is arbitrage free and complete. Furthermore, there exist a uniformly bounded $\theta(\cdot) = \sigma(t)^{-1}B(t)$. Note that $\theta(\cdot)$ is the only process satisfying $\sigma(\cdot)\theta(\cdot) = B(\cdot)$, which is called the *risk premium process.*

As in Chapter 2, define $\rho(\cdot)$ to be the deflator process as

$$\begin{cases} d\rho(t) = \rho(t)[-r(t)dt - \theta(t)'dW(t)], \\ \rho(0) = 1, \end{cases} \tag{3.2}$$

or, equivalently,

$$\rho(t) = \exp\left\{-\int_0^t [r(s) + \frac{1}{2}|\theta(s)|^2]ds - \int_0^t \theta(s)'dW(s)\right\}. \tag{3.3}$$

Then we have

$$x(t) = \rho(t)^{-1}E(\rho(T)x(T)|\mathcal{F}_t), \ \ \forall t \in [0,T] \tag{3.4}$$

for any wealth process $x(\cdot)$.

With (3.4), the wealth process $x(\cdot)$ is nonnegative if and only if the terminal wealth $x(T)$ is nonnegative. From the economic standpoint, this is a consequence of the fact that there exists a risk neutral probability measure under which the *discounted* wealth process is a martingale. Hence if the terminal wealth is nonnegative, then so is the discounted wealth process and thus $x(\cdot)$. This property can help us greatly simplify our problem, which we formulate a little later.

It should be emphasized an important point concerning the way we specify our trading strategies. Most papers in the research literature define a trading strategy or portfolio, say $u(\cdot)$, as the (vector of) proportions or fractions of wealth allocated to different assets, perhaps with some other "technical" constraints such as $\int_0^T |u(t)|^2 dt < \infty$, a.s., being specified (see, e.g., Cvitanic and Karatzas (1992) and Karatzas and Shreve (1998)). With this definition, and if additionally the self-financing property is postulated, then the wealth at any time $t \geq 0$ can be shown to be proportional to the wealth at time $t = 0$, in the sense that $x(t) = x_0\tilde{x}(t),$ where $\tilde{x}(t)$ is an (almost surely) strictly positive process. In fact, with a proportional, self-financing strategy $u(\cdot)$ satisfying the above condition, it can be shown that the wealth process is a unique strong solution of the following equation

$$\begin{cases} dx(t) = x(t)[r(t) + B(t)'u(t)]dt + x(t)u(t)'\sigma(t)dW(t), \\ x(0) = x_0. \end{cases} \tag{3.5}$$

Thus, $x(t) = x_0\tilde{x}(t),$ where

$$\tilde{x}(t) = \exp\Big\{ \int_0^t \Big([r(s) + B(s)'u(s)]^2 - \frac{1}{2}|u(s)'\sigma(s)|^2\Big) ds + \int_0^t u(s)'\sigma(s)dW(s)\Big\}.$$

Consequently, with proportional, self-financing strategies satisfying the above condition, the wealth process is *strictly* positive if the initial wealth x_0 is strictly positive. In fact, in this case the value $x = 0$ becomes a natural barrier of the wealth process.

However, in our model, with the portfolio defined to be the amounts of money allocated to different assets, the wealth process can take zero or negative values, and we require the nonnegativity of the wealth as an *additional constraint* rather than as a by-product of the "proportions of wealth" approach. Clearly the class of admissible, proportional, self-financing strategies is a proper sub-class of our set of admissible self-financing strategies. In fact, any admissible strategy $\pi(\cdot)$ which produces a (strictly) positive wealth process

$x(t) > 0$ gives rise to a proportional strategy, defined as $u(t) := \frac{\pi(t)}{x(t)}$. On the other hand, any proportional strategy $u(\cdot)$ gives rise to a "monetary amount" strategy $\pi(\cdot)$ defined as $\pi(t) = u(t)x(t)$. We will see later that our final solutions involve strategies that can*not* be expressed as proportional ones. Thus our model is *fundamentally different* from approaches based upon (2.5).

Now let us formulate the no bankruptcy mean-variance portfolio selection problem.

Definition 3.2.1 Consider the following optimization problem parameterized by $z \in \mathbf{R}$:

$$\begin{array}{ll} \text{Minimize} & \text{Var}\ x(T) \equiv Ex(T)^2 - z^2, \\ \text{subject to} & \begin{cases} Ex(T) = z, \\ x(T) \geq 0, \quad \text{a.s.,} \\ \pi(\cdot) \in L^2_{\mathcal{F}}(0, T; \mathbf{R}^m), \\ (x(\cdot), \pi(\cdot)) \text{ satisfies equation (2.5).} \end{cases} \end{array} \tag{3.6}$$

The optimal portfolio for this problem (corresponding to a fixed z) is called a **variance minimizing** portfolio, and the set of all points $(\text{Var}\ x^*(T), z)$, where $\text{Var}\ x^*(T)$ denotes the optimal value of (3.6) corresponding to z and z runs over $\mathbf{R}$, is called the **variance minimizing frontier**.

The efficient frontier, to be defined in Section 3.6, is a portion of the minimizing variance frontier. Once the minimizing variance frontier is identified, the efficient frontier can be easily obtained as an appropriate subset of the former[3]; see Section 3.6. Hence in this chapter we shall focus on problem (3.6).

If the initial wealth x_0 of the agent is zero and if the constraint $x(T) \geq 0$ is in force, then it follows from (3.4) that $x(t) \equiv 0$ under all admissible $\pi(\cdot)$. On the other hand, if z is set to be 0, then the constraints of (3.6) yield $x(T) = 0$, a.s., which in turn leads to $x(t) \equiv 0$ by (3.4). Hence to eliminate these trivial cases from consideration we assume from now on that

$$x_0 > 0, \quad z > 0. \tag{3.7}$$

[3]In some of the literature, problem (3.6) itself is defined as the mean–variance portfolio selection problem, with z required to be in a certain range.

To solve problem (3.6) we use an extension of the risk-neutral computational approach that was first introduced by Pliska [52], [53]. The idea is to decompose the problem into two sub-problems, the first of which is to find the optimal attainable wealth X^*, that is, the random variable that is the optimal value of all possible $x(T)$ obtainable by admissible portfolios. The second sub-problem is to find the trading strategy $\pi(\cdot)$ that replicates X^*, which is essentially a martingale representation problem.

To be specific, the first sub-problem is

$$\begin{array}{ll} \text{Minimize} & EX^2 - z^2, \\ \text{subject to} & \begin{cases} EX = z, \\ E[\rho(T)X] = x_0, \\ X \in L^2(\mathcal{F}_T, \mathbf{R}), \quad X \geq 0, \text{ a.s..} \end{cases} \end{array} \tag{3.8}$$

Assuming a solution X^* exists for this problem, consider the following terminal-valued equation:

$$\begin{cases} dx(t) = [r(t)x(t) + B(t)\pi(t)]dt + \pi(t)'\sigma(t)dW(t), \\ x(T) = X^*. \end{cases} \tag{3.9}$$

The following result verifies that problems (3.8) and (3.9) indeed lead to a solution of our original problem.

Theorem 3.2.1 *If $(x^*(\cdot), \pi^*(\cdot))$ is optimal for problem (3.6), then $x^*(T)$ is optimal for problem (3.8) and $(x^*(\cdot), \pi^*(\cdot))$ satisfies (3.9). Conversely, if X^* is optimal for problem (3.8), then (3.9) must have a solution $(x^*(\cdot), \pi^*(\cdot))$ which is an optimal solution for (3.6).*

Proof. Suppose $(x^*(\cdot), \pi^*(\cdot))$ is optimal for problem (3.6). First of all, by virtue of (3.4) we have $E[\rho(T)x^*(T)] = x_0$. Hence $x^*(T)$ is feasible for problem (3.8). Assume there is another feasible solution, denoted by Y, of problem (3.8) with

$$EY^2 < Ex^*(T)^2. \tag{3.10}$$

The following linear BSDE

$$\begin{cases} dx(t) = [r(t)x(t) + z(t)'\theta(t)]dt + z(t)'dW(t), \\ x(T) = Y \end{cases} \tag{3.11}$$

admits a unique square-integrable, $\mathcal{F}_t$-adapted solution $(x(\cdot), z(\cdot))$ since the coefficients of (3.11) are uniformly bounded due to the underlying assumptions. Write $\pi(t) = (\sigma(t)')^{-1}z(t)$, which is square integrable due to the uniform boundedness of $(\sigma(t)')^{-1}$. Hence $\pi(\cdot)$ is an admissible portfolio, and $(x(\cdot), \pi(\cdot))$ satisfies the same dynamics of (2.2.5). Moreover, it follows from (3.4) that

$$x(0) = E[\rho(T)Y] = x_0,$$

where the second equality is due to the feasibility of Y to problem (3.8). This implies $(x(\cdot), \pi(\cdot))$ is a feasible solution to (3.6). However, (3.10) yields $Ex(T)^2 = EY^2 < Ex^*(T)^2$, contradicting the optimality of $(x^*(\cdot), \pi^*(\cdot))$.

Conversely, let X^* be optimal for problem (3.8). Then by a similar argument to that above, and using the BSDE (3.11) with terminal condition $x(T) = X^*$, one sees that one can construct a feasible solution $(x^*(\cdot), \pi^*(\cdot))$ to (3.6). Moreover, if there is another feasible solution $(x(\cdot), \pi(\cdot))$ to (3.6) that is better than $(x^*(\cdot), \pi^*(\cdot))$, then $x(T)$ would be better than X^* for problem (3.8), leading to a contradiction. □

Remark 3.2.1 By virtue of the above theorem, solving the variance minimizing problem boils down to solving the optimization problem (3.8). Once (3.8) is solved, the solution to (3.9) can be obtained via standard BSDE theory.

3.3 Feasibility

Since problem (3.6) involves several constraints, the first issue is its feasibility, which is the subject of this section.

Proposition 3.3.1 *Problem (3.6) either has no feasible solution or it admits a unique optimal solution.*

Proof: In view of Remark 3.2.1 it suffices to investigate the feasibility of (3.8). Now (3.8) can be regarded as an optimization problem on the Hilbert space $L^2(\mathcal{F}_T, R)$, with the constraint set

$$D := \{Y \in L^2(\mathcal{F}_T, \mathbf{R}) : \ EY = z, E[\rho(T)Y] = x_0, Y \geq 0\}.$$

If D is nonempty, say with $Y_0 \in D$, then an optimal solution of (3.8), if any, must be in the set $D' := D \cap \{EY^2 \leq EY_0^2\}$. In this case, clearly D' is a nonempty, bounded, closed convex set in $L^2(\mathcal{F}_T, \mathbf{R})$. Moreover, the cost functional of (3.8) is strictly convex on D' with a lower bound $-z^2$. Hence (3.8) must admit a unique optimal solution. $\Box$

Define

$$\begin{aligned} a &:= \inf_{Y \in L^2(\mathcal{F}_T,\mathbf{R}), Y\geq 0, EY>0} \tfrac{E[\rho(T)Y]}{EY}, \\ b &:= \sup_{Y \in L^2(\mathcal{F}_T,\mathbf{R}), Y\geq 0, EY>0} \tfrac{E[\rho(T)Y]}{EY}. \end{aligned} \tag{3.12}$$

As will be evident from the sequel, the values a and b are critical. The following representations of a and b are useful.

Proposition 3.3.2 *We have the following representation*

$$\begin{aligned} a &= \inf\{\eta \in \mathbf{R} : P(\rho(T) < \eta) > 0\}, \\ b &= \sup\{\eta \in \mathbf{R} : P(\rho(T) > \eta) > 0\}. \end{aligned} \tag{3.13}$$

Proof: Denote $\bar{a} := \inf\{\eta \in \mathbf{R} : P(\rho(T) < \eta) > 0\}$. For any η satisfying $P(\rho(T) < \eta) > 0$, take $Y := \mathbf{1}_{\rho(T)<\eta}$. Then

$$Y \in L^2(\mathcal{F}_T, \mathbf{R}), \;\; Y \geq 0, \;\; EY > 0, \;\; \text{and} \;\; \frac{E[\rho(T)Y]}{EY} < \eta.$$

As a result, by the definition of a, we have $a \leq \frac{E[\rho(T)Y]}{EY} < \eta$. Hence $a \leq \bar{a}$. Conversely, by the definition of $\bar{a}$ we must have $P(\rho(T) < \bar{a} - \epsilon) = 0$ for any $\epsilon > 0$, namely, $\rho(T) \geq \bar{a} - \epsilon$, a.s.. Hence for any $Y \in L^2(\mathcal{F}_T, \mathbf{R})$ with $Y \geq 0, EY > 0$, we have $\frac{E[\rho(T)Y]}{EY} \geq \bar{a} - \epsilon$. This implies $a \geq \bar{a} - \epsilon$ for any $\epsilon > 0$; thus $a \geq \bar{a}$.

We have now proved the first equality of (3.13). The second one can be proved in a similar fashion. $\Box$

Remark 3.3.1 When the risk premium process $\theta(\cdot)$ is deterministic, and when $\int_0^T |\theta(t)|^2 dt > 0$, the exponent in (3.3) at $t = T$ is the sum of a bounded random variable and a normal random variable with a strictly positive variance; hence $a = 0, b = +\infty$ by Proposition 3.3.2. But when $\theta(\cdot)$ is a stochastic process, both $a > 0$ and $b < +\infty$ are possible even if $\int_0^T |\theta(t)|^2 dt > 0$, a.s.. To show this, by (3.3) it suffices to construct an example where $\int_0^T \theta(t) dW(t)$ is uniformly bounded. Indeed, consider a market with one bank account and one stock with the corresponding one-dimensional standard Brownian Motion $W(t)$. For a given real number $K > 0$, define

$$\tau := \begin{cases} \inf\{t \geq 0 : |W(t)| > K\}, & \text{if } \sup_{0\leq t\leq T} |W(t)| > K, \\ T, & \text{if } \sup_{0\leq t\leq T} |W(t)| \leq K. \end{cases} \tag{3.14}$$

Take $r(t) = 0.1$, $b(t) = 0.1 + \mathbf{1}_{t\le\tau}$ and $\sigma(t) = 1$. Thus $\theta(t) = \mathbf{1}_{t\le\tau}$. Then $\int_0^T \theta(t)dW(t) = W(\tau)$, which is uniformly bounded by K.

The next result is very important, for it specifies an interval such that our problem (3.6) has a solution almost if and only if the desired expected wealth z takes a value in this interval.

Proposition 3.3.3 *If $a < \frac{x_0}{z} < b$, then there must be a feasible solution to (3.6). Conversely, if (3.6) has a feasible solution, then it must be that $a \le \frac{x_0}{z} \le b$.*

Proof: Assume $a < \frac{x_0}{z} < b$. Again we only need to show the feasibility of the problem (3.8). By the definition of a and b, for any $x_0 > 0$ and $z > 0$ with $a < \frac{x_0}{z} < b$ there exist $Y_1, Y_2 \in \{Y \in L^2(\mathcal{F}_T, \mathbf{R}) : \ Y \ge 0, EY > 0\}$ such that

$$\frac{E[\rho(T)Y_1]}{EY_1} < \frac{x_0}{z} < \frac{E[\rho(T)Y_2]}{EY_2}.$$

Define a function

$$f(\lambda) := \frac{E[\rho(T)(\lambda Y_1 + (1-\lambda)Y_2)]}{E[\lambda Y_1 + (1-\lambda)Y_2]} = \frac{\lambda E[\rho(T)Y_1] + (1-\lambda)E[\rho(T)Y_2]}{\lambda EY_1 + (1-\lambda)EY_2}, \quad \lambda \in [0,1].$$

Then f is continuous on $[0,1]$ with $f(1) < \frac{x_0}{z} < f(0)$, so there exists a $\lambda_0 \in (0,1)$ such that $\frac{x_0}{z} = f(\lambda_0) = \frac{E[\rho(T)(\lambda_0 Y_1 + (1-\lambda_0)Y_2)]}{E[\lambda_0 Y_1 + (1-\lambda_0)Y_2]}$. Set $Y_0 := \lambda_0 Y_1 + (1-\lambda_0)Y_2$ and $Y^* := zY_0/E[Y_0]$. Then clearly $Y^* \in L^2(\mathcal{F}_T, \mathbf{R})$, $Y^* \ge 0$, $E(Y^*) = z$, and

$$E[\rho(T)Y^*] = zf(\lambda_0) = x_0.$$

This shows that Y^* is a feasible solution of (3.8).

Conversely, if there is a feasible solution of (3.6), then (3.8) also has a feasible solution, say Y^*. Hence $Y^* \in L^2(\mathcal{F}_T, \mathbf{R})$, $Y^* \ge 0$, and $E[Y^*] = z$. Thus,

$$\frac{x_0}{z} = \frac{E[\rho(T)Y^*]}{EY^*} \ge a.$$

Similarly, $\frac{x_0}{z} \le b$. □

One naturally wonders what can be said about the feasibility of (3.6) when $\frac{x_0}{z} = a$ or b. The answer is that at these "boundary" points, (3.6) may *or* may not be feasible, as can be seen from the following example.

Example 3.3.1 First consider the process $\theta(\cdot)$ as given in Remark 3.1, namely $\theta(t) = \mathbf{1}_{t\le\tau}$, where τ is defined by (3.14) for a one-dimensional standard Brownian motion $W(t)$ and a given real number $K > 0$. Let $r(t) := -\frac{|\theta(t)|^2}{2}$. Then it follows from (3.3) that $\rho(T) = e^{-\int_0^T \theta(t)dW(t)} = e^{-W(\tau)}$. Now

$$a = \inf\{\eta \in \mathbf{R} : P(\rho(T) < \eta) > 0\} = e^{-K},$$

whereas

$$P(\rho(T) = a) = P(W(\tau) = K) = 1 - P(\sup_{0\le t\le T} |W(t)| < K) > 0.$$

Take $Y := \mathbf{1}_{\rho(T)=a}$. Then $Y \ge 0$, $EY > 0$ and $E[\rho(T)Y] = aP(\rho(T) = a)$. Hence with $x_0 := aP(\rho(T) = a) > 0$ and $z := EY = P(\rho(T) = a) > 0$, we have $\frac{x_0}{z} = a$ while Y is a feasible solution to (3.8).

Next, let $\theta(\cdot)$ be the same as above, and $r(t) = -\frac{|\theta(t)|^2}{2} - \mathbf{1}_{t<\tau}$. Then $\rho(T) = e^{-W(\tau)+\tau}$, $a = \inf\{\eta \in \mathbf{R} : P(\rho(T) < \eta) > 0\} = e^{-K}$, and

$$P(\rho(T) > a) \ge P(W(\tau) \le K) = 1. \tag{3.15}$$

If there is a feasible solution Y to (3.8) for certain $x_0 > 0$ and $z > 0$ with $\frac{x_0}{z} = a$, or $\frac{E[\rho(T)Y]}{EY} = a$, then

$$E[(\rho(T) - a)Y] = 0,$$

implying $Y = 0$ a.s. in view of (3.15). Thus, $EY = 0$ leading to a contradiction. So (3.8) has no feasible solution when $\frac{x_0}{z} = a$.

We summarize most of the results in this section as follows:

Theorem 3.3.1 *If $a < \frac{x_0}{z} < b$, then the minimizing variance problem (3.6) is feasible and must admit a unique optimal solution. In particular, if the process $\theta(\cdot)$ is deterministic with $\int_0^T |\theta(t)|^2 dt > 0$, then (3.6) must have a unique optimal solution for any $x_0 > 0, z > 0$.*

3.4 Solution to (3.8): the optimal attainable wealth

In this section we present the complete solution to the auxiliary problem (3.8). First a preliminary result involving Lagrange multipliers follows.

Proposition 3.4.1 *Let $D \subset L^2(\mathcal{F}_T, \mathbf{R})$ be a convex set, $a_i \in \mathbf{R}$, and $\xi_i \in L^2(\mathcal{F}_T, \mathbf{R})$, $i = 1, 2, \cdots, l$, be given, and let f be a scalar-valued convex function on $\mathbf{R}$. If the problem*

$$\begin{array}{ll} \text{minimize} & Ef(Y), \\ \text{subject to} & \begin{cases} E[\xi_i Y] = a_i, & i = 1, 2, \cdots, l, \\ Y \in D \end{cases} \end{array} \tag{3.16}$$

has a solution Y^*, then there exists an l-dimensional deterministic vector $(\lambda_1, \cdots, \lambda_l)'$ such that Y^* also solves the following

$$\begin{array}{ll} \text{minimize} & E[f(Y) - Y \sum_{i=1}^{l} \lambda_i \xi_i], \\ \text{subject to} & Y \in D. \end{array} \tag{3.17}$$

Conversely, if Y^* solves (3.17) for some $(\lambda_1, \cdots, \lambda_l)'$, then it must also solve (3.16) with $a_i = E[\xi_i Y^*]$.

Proof: Let Y^* solve (3.16). Define a set $\Delta := \{(E[\xi_1 Y], \cdots, E[\xi_l Y])' : Y \in D\} \subseteq \mathbf{R}^l$, which is clearly a convex set, and a function

$$g(x) \equiv g(x_1, \cdots, x_l) := \inf_{E[\xi_i Y] = x_i, i=1,\cdots,l, Y \in D} E[f(Y)], \quad x \in \Delta.$$

In view of the assumptions, g is a convex function on Δ. By the convex separation theorem, for the given $a = (a_1, \cdots, a_l)'$, there exists an l-dimensional vector $\lambda = (\lambda_1, \cdots, \lambda_l)'$ such that $g(x) \geq g(a) + \lambda'(x - a)$, $\forall x \in \Delta$. Equivalently, $g(x) - \lambda' x \geq g(a) - \lambda' a$. Now, for any $Y \in D$,

$$\begin{aligned} E[f(Y) - Y \textstyle\sum_{i=1}^{l} \lambda_i \xi_i] & \geq g(E[\xi_1 Y], \cdots, E[\xi_l Y]) - \textstyle\sum_{i=1}^{l} \lambda_i E[\xi_i Y] \\ & \geq g(a) - \lambda' a \\ & = E[f(Y^*) - Y^* \textstyle\sum_{i=1}^{l} \lambda_i \xi_i], \end{aligned}$$

implying that Y^* solves (3.17).

Conversely, if Y^* solve (3.17), then for any $Y \in D$ satisfying $E[\xi_i Y] = E[\xi_i Y^*]$, we have

$$E[f(Y^*) - Y^* \sum_{i=1}^{l} \lambda_i \xi_i] \leq E[f(Y) - Y \sum_{i=1}^{l} \lambda_i \xi_i] = E[f(Y) - Y^* \sum_{i=1}^{l} \lambda_i \xi_i].$$

Hence $E[f(Y^*)] \leq E[f(Y)]$, thereby proving the desired result. □

We now solve problem (3.8) by using Proposition 3.4.1 to transform it to an equivalent problem that has two Lagrange multipliers and only one constraint: $X \geq 0$.

Theorem 3.4.1 *If problem (3.8) admits a solution X^*, then $X^* = (\lambda - \mu\rho(T))^+$, where the pair of scalars (λ, μ) solves the system of equations*

$$\begin{cases} E[(\lambda - \mu\rho(T))^+] = z, \\ E[\rho(T)(\lambda - \mu\rho(T))^+] = x_0. \end{cases} \tag{3.18}$$

Conversely, if (λ, μ) satisfies (3.18), then $X^* := (\lambda - \mu\rho(T))^+$ must be an optimal solution of (3.8).

Proof: If X^* solves problem (3.8), then by Proposition 3.4.1 there exists a pair of constants $(2\lambda, -2\mu)$ such that X^* also solves

$$\begin{aligned} &\text{minimize} \quad E[X^2 - 2\lambda X + 2\mu\rho(T)X] - z^2, \\ &\text{subject to} \quad X \geq 0, \ \text{ a.s.}. \end{aligned} \tag{3.19}$$

However, the objective function of (3.19) equals

$$E[X - (\lambda - \mu\rho(T))]^2 - z^2 - E[\lambda - \mu\rho(T)]^2.$$

Hence problem (3.19) has an obvious unique solution $(\lambda - \mu\rho(T))^+$ which must then coincide with X^*. In this case, the two equations in (3.18) are nothing else than the two equality constraints in problem (3.8).

The converse result of the theorem can be proved similarly in view of Proposition 3.4.1 □

Observe that if the non-negativity constraint $X \geq 0$ is removed from problem (3.8), then the optimal solution to such a relaxed problem is simply $X^* = \lambda - \mu\rho(T)$, with the constants λ and μ satisfying

$$\begin{cases} E[\lambda - \mu\rho(T)] = z, \\ E[\rho(T)(\lambda - \mu\rho(T))] = x_0. \end{cases} \tag{3.20}$$

Since these equations are linear, the solution is immediate:

$$\lambda = \frac{zE[\rho(T)^2] - x_0 E[\rho(T)]}{\operatorname{Var}\rho(T)}, \quad \mu = \frac{zE[\rho(T)] - x_0}{\operatorname{Var}\rho(T)}.$$

But for problem (3.8) the existence and uniqueness of Lagrange multipliers λ and μ satisfying (3.18) is a more delicate issue, which we discuss in the following section.

3.5 Existence and uniqueness of Largrange multipliers

By virtue of Theorem 3.4.1, an optimal solution to (3.8) is obtained explicitly if the system of equations (3.18) for Lagrange multipliers admits solutions. In this section we study the unique solvability of (3.18). For notational simplicity we rewrite (3.18) as

$$\begin{cases} E[(\lambda - \mu Z)^+] = z, \\ E[(\lambda - \mu Z)^+ Z] = x_0, \end{cases} \tag{3.21}$$

where $Z := \rho(T)$. First we have three preliminary lemmas.

Lemma 3.5.1 *For any random variable X and real number c,*

$$E[X(c-X)] - E[X]E[c-X] \le 0, \qquad E[X(X-c)] - E[X]E[X-c] \ge 0.$$

Proof: We have

$$\begin{aligned} E[X(c-X)] - E[X]E[c-X] &= -E[X^2] + (EX)^2 \le 0, \\ E[X(X-c)] - E[X]E[X-c] &= E[X^2] - (EX)^2 \ge 0. \end{aligned}$$

□

Lemma 3.5.2 *The function $R_1(\eta) := \frac{E[(\eta-Z)^+Z]}{E[(\eta-Z)^+]}$ is continuous and strictly increasing for $\eta \in (a, +\infty)$, and the function $R_2(\eta) := \frac{E[(Z-\eta)^+Z]}{E[(Z-\eta)^+]}$ is continuous and strictly decreasing for $\eta \in (-\infty, b)$, where a and b are given in (3.13).*

Proof: Let us first observe that in view of characterization (3.13) we have that $P(Z < \eta) > 0$ for any $\eta > a$, and that $P(Z > \eta) > 0$ for any $\eta < b$. Consequently, $P((\eta - Z)^+ > 0) > 0$ for any $\eta > a$ and $P((Z - \eta)^+ > 0) > 0$ for any $\eta < b$. Thus the following inequalities are satisfied: $E[(\eta - Z)^+] > 0$ for $\eta > a$, and $E[(Z - \eta)^+] > 0$ for $\eta < b$. This verifies continuity of both functions.

To prove the strict monotonicity of $R_1(\cdot)$, take any $\eta_1 > \eta_2 > a$. Then we have

$$\begin{aligned} \frac{E[(\eta_2-Z)^+Z]}{E[(\eta_2-Z)^+]} &= \frac{E[(\eta_2-Z)Z|Z<\eta_2]}{E[(\eta_2-Z)|Z<\eta_2]} \\ &\le E(Z|Z<\eta_2) \quad \text{(by Lemma 3.5.1)} \\ &= \frac{E[Z\mathbf{1}_{Z<\eta_2}]}{E[\mathbf{1}_{Z<\eta_2}]} \\ &< \eta_2 \\ &\le \frac{E[(\eta_1-Z)Z\mathbf{1}_{\eta_2\le Z\le\eta_1}]}{E[(\eta_1-Z)\mathbf{1}_{\eta_2\le Z\le\eta_1}]}. \end{aligned} \tag{3.22}$$

Note that in particular the above inequalities imply that

$$\frac{E[(\eta_1-\eta_2)Z\mathbf{1}_{Z<\eta_2}]}{E[(\eta_1-\eta_2)\mathbf{1}_{Z<\eta_2}]} = \frac{E[Z\mathbf{1}_{Z<\eta_2}]}{E[\mathbf{1}_{Z<\eta_2}]} < \frac{E[(\eta_1-Z)Z\mathbf{1}_{\eta_2\le Z\le\eta_1}]}{E[(\eta_1-Z)\mathbf{1}_{\eta_2\le Z\le\eta_1}]}. \tag{3.23}$$

On the other hand,

$$\begin{aligned}\frac{E\{[(\eta_1-Z)^+-(\eta_2-Z)^+]Z\}}{E[(\eta_1-Z)^+-(\eta_2-Z)^+]} &= \frac{E[(\eta_1-\eta_2)Z\mathbf{1}_{Z<\eta_2}]+E[(\eta_1-Z)Z\mathbf{1}_{\eta_2\le Z\le\eta_1}]}{E[(\eta_1-\eta_2)\mathbf{1}_{Z<\eta_2}]+E[(\eta_1-Z)\mathbf{1}_{\eta_2\le Z\le\eta_1}]}\\ &> \frac{E[(\eta_1-\eta_2)Z\mathbf{1}_{Z<\eta_2}]}{E[(\eta_1-\eta_2)\mathbf{1}_{Z<\eta_2}]}\\ &\ge \frac{E[(\eta_2-Z)^+Z]}{E[(\eta_2-Z)^+]},\end{aligned} \tag{3.24}$$

where the first inequality is due to (3.23) and the familiar inequality

$$\frac{x_1+x_2}{y_1+y_2} > \frac{x_1}{y_1} \qquad \text{if} \quad \frac{x_2}{y_2} > \frac{x_1}{y_1} \quad \text{and} \quad y_1, y_2 > 0, \tag{3.25}$$

and the last inequality follows from (3.22). Finally,

$$\begin{aligned}\frac{E[(\eta_1-Z)^+Z]}{E[(\eta_1-Z)^+]} &= \frac{E[(\eta_2-Z)^+Z]+E\{[(\eta_1-Z)^+-(\eta_2-Z)^+]Z\}}{E[(\eta_2-Z)^+]+E[(\eta_1-Z)^+-(\eta_2-Z)^+]}\\ &> \frac{E[(\eta_2-Z)^+Z]}{E[(\eta_2-Z)^+]},\end{aligned} \tag{3.26}$$

owing to (3.24) and inequality (3.25). This shows that $R_1(\cdot)$ is strictly increasing. Similarly, we can prove that $R_2(\cdot)$ is strictly decreasing. □

Lemma 3.5.3 *We have the following interval representations of the respective sets:*

$$\{R_1(\eta) : \eta > a\} = (a, E[Z]), \tag{3.27}$$

$$\{R_2(\eta) : \eta < 0\} = (E[Z], \frac{E[Z^2]}{E[Z]}), \tag{3.28}$$

$$\{R_2(\eta) : 0 \le \eta < b\} = [\frac{E[Z^2]}{E[Z]}, b). \tag{3.29}$$

Proof: By the definition of a we have $P(Z < a) = 0$. In other words $Z \ge a$, a.s. Hence

$$R_1(\eta) = \frac{E[(\eta-Z)^+Z]}{E[(\eta-Z)^+]} \ge a, \ \ \forall \eta > a. \tag{3.30}$$

Meanwhile,

$$E[(\eta-Z)^+Z] \le E[(\eta-Z)^+\eta] = \eta E[(\eta-Z)^+],$$

leading to

$$R_1(\eta) \le \eta, \ \ \forall \eta > a. \tag{3.31}$$

Combining (3.30) and (3.31) we conclude

$$\lim_{\eta \to a+} R_1(\eta) = a. \tag{3.32}$$

On the other hand,

$$\begin{aligned} \lim_{\eta \to +\infty} R_1(\eta) &= \lim_{\eta \to +\infty} \tfrac{E[(\eta - Z)^+ Z]}{E[(\eta - Z)^+]} \\ &= \lim_{\eta \to +\infty} \tfrac{E[(1 - Z/\eta)^+ Z]}{E[(1 - Z/\eta)^+]} \\ &= E[Z]. \end{aligned} \tag{3.33}$$

Hence, (3.27) follows from the fact that $R_1(\eta)$ is continuous and strictly increasing.

Next, observe that since Z is almost surely positive, then for every $\eta \leq 0$ we have that $E[(Z-\eta)^+ Z] = E[(Z-\eta)Z]$ and $E[(Z-\eta)^+] = E(Z-\eta)$. Consequently, we obtain that

$$\lim_{\eta \to -\infty} R_2(\eta) = \lim_{\eta \to -\infty} \frac{E[(Z-\eta)^+ Z]}{E[(Z-\eta)^+]} = \lim_{\eta \to -\infty} \frac{E[Z^2] - \eta E[Z]}{E[Z] - \eta} = E[Z],$$

and

$$R_2(0) = \frac{E[Z^2]}{E[Z]}.$$

The above as well as the strict monotonicity of $R_2(\cdot)$ imply (3.28). Finally, an argument analogous to the one that lead to (3.32) yields

$$\lim_{\eta \to b-} R_2(\eta) = b,$$

and this implies (3.29). □

Now we are in a position to present our main results on the unique solvability of equations (3.21). In particular, we characterize the signs of the two Lagrange multipliers.

Theorem 3.5.1 *Equations (3.21) have a unique solution* (λ, μ) *for any* $x_0 > 0, z > 0$ *satisfying* $a < \frac{x_0}{z} < b$. *Moreover,*

(1) $\lambda = z, \mu = 0$ *if* $\frac{x_0}{z} = E[Z]$;

(2) $\lambda > 0, \mu > 0$ *if* $a < \frac{x_0}{z} < E[Z]$;

(3) $\lambda \leq 0, \mu < 0$ *if* $\frac{E[Z^2]}{E[Z]} \leq \frac{x_0}{z} < b$;

(4) $\lambda > 0, \mu < 0$ *if* $E[Z] < \frac{x_0}{z} < \frac{E[Z^2]}{E[Z]}$.

Proof: First of all, if $EZ^2 = (EZ)^2$, then the variance of Z is zero or Z is a deterministic constant almost surely. Hence $a = b$ by (3.12), which violates the assumption of the theorem. Consequently, $E[Z^2] > (E[Z])^2$. On the other hand, again by (3.12) we have immediately (by letting $Y = 1$ and $Y = Z$ in $\frac{E[ZY]}{E[Y]}$, respectively)

$$a \leq E[Z] < \frac{E[Z^2]}{E[Z]} \leq b,$$

where it is important to note the strict inequality above.

We now examine the four cases. Case (1) is easy, for when $\frac{x_0}{z} = E[Z]$, one directly verifies that $\lambda = z$, $\mu = 0$ solve (3.21).

For the other three cases we must have $\mu^* \neq 0$ for any solution (λ^*, μ^*) of (3.21), for otherwise in view of (3.21) we have $\lambda^* = z$ and $\lambda^* E[Z] = x_0$ leading to $\frac{x_0}{z} = E[Z]$ which is Case (1).

Next, observe that if $\mu^* > 0$, then $(\eta, \mu) := (\frac{\lambda^*}{\mu^*}, \mu^*)$ is a solution of the following equations

$$\begin{cases} \frac{E[(\eta - Z)^+ Z]}{E[(\eta - Z)^+]} = \frac{x_0}{z}, \\ E[(\eta - Z)^+] = \frac{z}{\mu}. \end{cases} \tag{3.34}$$

Likewise, if $\mu^* < 0$, then $(\eta, \mu) := (\frac{\lambda^*}{\mu^*}, \mu^*)$ is a solution of the following equations

$$\begin{cases} \frac{E[(Z - \eta)^+ Z]}{E[(Z - \eta)^+]} = \frac{x_0}{z} \\ E[(Z - \eta)^+] = -\frac{z}{\mu}. \end{cases} \tag{3.35}$$

Now for case (2) where $a < \frac{x_0}{z} < E[Z]$ it follows from Lemma 3.5.3 that the first equation of (3.34) admits a unique solution $\eta^* > a \geq 0$ and (3.35) admits no solution. Set

$$\mu^* := \frac{z}{E[(\eta^* - Z)^+]} > 0, \quad \lambda^* := \eta^* \mu^* > 0.$$

Then (λ^*, μ^*) is the unique solution for (3.21).

If $\frac{E[Z^2]}{E[Z]} \leq \frac{x_0}{z} < b$, which is case (3), then by Lemma 3.5.3 the first equation of (3.35) admits a unique solution $\eta^* \geq 0$ and (3.34) admits no solution. Set

$$\mu^* := -\frac{z}{E[(Z - \eta^*)^+]} < 0, \quad \lambda^* := \eta^* \mu^* \leq 0.$$

Then (λ^*, μ^*) is the unique solution for (3.21).

Finally, in case (4) where $E[Z] < \frac{x_0}{z} < \frac{E[Z^2]}{E[Z]}$, Lemma 3.5.3 yields that the first equation of (3.35) admits a unique solution $\eta^* < 0$ and (3.34) admits no solution. Letting

$$\mu^* := -\frac{z}{E[(Z - \eta^*)^+]} < 0, \quad \lambda^* := \eta^* \mu^* > 0,$$

we get that (λ^*, μ^*) uniquely solves (3.21). $\square$

Remark 3.5.1 In Theorem 3.5.1, we separate the feasible set (a, b) by two critical points: EZ and $\frac{EZ^2}{EZ}$. It's easy to see that EZ is the reciprocal of the expected return rate of risk-free investment, while $\frac{EZ^2}{EZ}$ is the reciprocal of the expected return rate of the investment on a special contingent claim $Z := \rho(T)$.

Observe that the Lagrange multipliers have a homogeneous property, for if one denotes by $(\lambda(x_0, z), \mu(x_0, z))$ the solution to (3.21) when taking $x_0 > 0$ and $z > 0$ as parameters, then clearly

$$\lambda(x_0, z) = x_0 \lambda(1, \frac{z}{x_0}), \quad \mu(x_0, z) = x_0 \mu(1, \frac{z}{x_0}).$$

In other words, the solution really depends only on the ratio z/x_0, which is essentially the expected return desired by the investor.

3.6 Efficient portfolios and efficient frontier

In this section we derive the efficient portfolios and efficient frontier of our mean–variance portfolio selection problem based on the variance minimizing portfolios and variance minimizing frontier. We fix the initial capital level $x_0 > 0$ for the rest of this section.

First we give the following definition, following p.6 in Markowitz [43].

Definition 3.6.1 The mean–variance portfolio selection problem with bankruptcy prohibition is formulated as the following multi-objective optimization problem

$$\begin{array}{ll} \text{Minimize} & (J_1(\pi(\cdot)), J_2(\pi(\cdot))) := (\mathrm{Var}\ x(T), -Ex(T)), \\ \text{subject to} & \begin{cases} x(T) \geq 0, \quad \text{a.s.}, \\ \pi(\cdot) \in L^2_{\mathcal{F}}(0, T; \mathbf{R}^m), \\ (x(\cdot), \pi(\cdot)) \text{ satisfies equation (2.5).} \end{cases} \end{array} \tag{3.36}$$

An admissible portfolio $\pi^*(\cdot)$ is called an *efficient portfolio* if there exists no admissible portfolio $\pi(\cdot)$ satisfying (3.36) such that

$$J_1(\pi(\cdot)) \leq J_1(\pi^*(\cdot)), \qquad J_2(\pi(\cdot)) \leq J_2(\pi^*(\cdot)), \tag{3.37}$$

with at least one of the inequalities holds strictly. In this case, we call $(J_1(\pi^*(\cdot)), -J_2(\pi^*(\cdot))) \in \mathbf{R}^2$ an *efficient point*. The set of all efficient points is called the *efficient frontier*.

In words, an efficient portfolio is one for which there does not exist another portfolio that has higher mean and no higher variance, and/or has less variance and no less mean at the terminal time T. In other words, an efficient portfolio is one that is *Pareto optimal*. The problem then is to identify *all* the efficient portfolios along with the efficient frontier.

By their very definitions the efficient frontier is a subset of the variance minimizing frontier, and efficient portfolios must be variance minimizing portfolios. In fact, an alternative definition of an efficient portfolio is the following. A variance minimizing portfolio $\tilde{\pi}_z$ corresponding to the terminal expected wealth z is called efficient if it is also *mean maximizing* in the following sense: $Ex^{\pi}(T) \leq Ex^{\tilde{\pi}_z}(T)$ for all portfolios π that satisfy the conditions

$$\begin{cases} \pi(\cdot) \in L^2_{\mathcal{F}}(0,T;\mathbf{R}^m), \\ (x^{\pi}(\cdot), \pi(\cdot)) \text{ satisfies equation (2.5)}, \\ x^{\pi}(T) \geq 0, \quad \text{a.s.}, \\ \text{Var } x^{\pi}(T) = \text{Var } x^{\tilde{\pi}_z}(T) \ , \end{cases} \tag{3.38}$$

where $x^{\pi}(\cdot)$ denotes the wealth process under a portfolio $\pi(\cdot)$ and with the initial wealth x_0.

The preceding discussion shows that our first task is to obtain variance minimizing portfolios, namely, the optimal trading strategies for problem (3.6).

Theorem 3.6.1 *The unique variance minimizing portfolio for (3.6) corresponding to $z > 0$, where $a < \frac{x_0}{z} < b$, is given by*

$$\pi^*(t) = (\sigma(t)')^{-1} z^*(t), \tag{3.39}$$

where $(x^*(\cdot), z^*(\cdot))$ is the unique solution to the BSDE

$$\begin{cases} dx(t) = [r(t)x(t) + z(t)'\theta(t)]dt + z(t)'dW(t) \\ x(T) = (\lambda - \mu\rho(T))^+, \end{cases} \tag{3.40}$$

with (λ, μ) being the solution to (3.18).

Proof: Since $\rho(\cdot)$ is the solution to (3.2), $\rho(T) \in L^2(\mathcal{F}_T, \mathbf{R})$. Meanwhile by Theorem 3.5.1 equation (3.18) admits a unique solution (λ, μ). By standard linear BSDE theory, (3.40) has a unique solution $(x^*(\cdot), z^*(\cdot)) \in L^2_{\mathcal{F}}(0,T;\mathbf{R}) \times L^2_{\mathcal{F}}(0,T;\mathbf{R}^m)$. Thus, the portfolio defined by (3.39) must be admissible. Now, the pair $(x^*(\cdot), \pi^*(\cdot))$ satisfies (3.9) with

$X^* = (\lambda - \mu\rho(T))^+$, the latter being the optimal solution of (3.8) by virtue of Theorem 3.4.1. Thus, Theorem 3.2.1 stipulates that $\pi^*(\cdot)$ must be optimal for (3.6). $\Box$

Theorem 3.6.1 asserts that a variance minimizing portfolio is the one that replicates the time-T payoff of the contingent claim $(\lambda - \mu\rho(T))^+$. Note that computing solutions of BSDE's like (3.40) is reasonably standard; see, for example, Ma, Protter, and Yong [39] or Ma and Yong [40]. In particular, if the market coefficients are deterministic, then it is possible to solve (3.40) explicitly via some partial differential equations; see Section 3.7 for details.

Our next result pinpoints the value of z corresponding to the riskless investment in our economy.

Theorem 3.6.2 *The variance minimizing portfolio corresponding to* $z = \frac{x_0}{E[\rho(T)]}$ *is a risk-free portfolio.*

Proof: By Theorem 3.5.1, $\lambda = z$ and $\mu = 0$ when $\frac{x_0}{z} = E[\rho(T)]$. The terminal wealth under the corresponding variance minimizing portfolio, say $\pi_0(\cdot)$, is therefore $x_0(T) = (\lambda - \mu\rho(T))^+ = \lambda = z$. Hence this portfolio is risk-free. $\Box$

In view of Theorem 3.6.2, the risk-free portfolio $\pi_0(\cdot)$ exists even when all the market parameters are random. Under $\pi_0(\cdot)$ a terminal payoff $\frac{x_0}{E[\rho(T)]}$ is guaranteed. Hence $E[\rho(T)]$ can be regarded as the risk-adjusted discount factor between 0 and T. We may explain this from another angle. Note in this case $x_0 = s_0 E_Q[S_0(T)^{-1}z]$, namely, the initial wealth x_0 is equal to the present value of a (sure) cash flow of z units at time $t = T$. Since our market is complete, there must be a portfolio having initial value x_0 and replicating this cash deterministic flow. Our portfolio $\pi_0(\cdot)$ is such a replicating portfolio. Note, however, that $\pi_0(\cdot)$ might involve exposure to the stocks. When the spot interest rates $r(t)$ are random, it is necessary to hedge the interest rate risk by taking a suitable position in the stocks; since the market is complete, this risk can be eliminated.

Due to the availability of the risk-free portfolio, it is sensible to restrict attention to values of the expected payoff satisfying $z \geq \frac{x_0}{E[\rho(T)]}$ when considering problem (3.6). On the other hand, by Proposition 3.3.3, z will be too large for the mean–variance problem to be feasible if $z > \frac{x_0}{a}$ ($\frac{x_0}{a}$ is defined to be ∞ if $a = 0$). Hence it is sensible to focus on values of the parameter z (the targeted mean terminal payoff) satisfying $\frac{x_0}{E[\rho(T)]} \leq z < \frac{x_0}{a}$ (In particular, in the special case where the interest rate process $r(\cdot)$ and the other parameters in the model are deterministic, then the relevant interval for the mean terminal payoff z is simply $[x_0 e^{\int_0^T r(t)dt}, \infty))$. For such values of z we then have the following economic interpretation of the optimal terminal wealth.

Proposition 3.6.1 *The unique variance minimizing portfolio for (3.6) corresponding to z with $\frac{x_0}{E[\rho(T)]} \leq z < \frac{x_0}{a}$ is a replicating portfolio for a European put option written on the fictitious asset $\mu\rho(\cdot)$ with a strike price $\lambda > 0$ and maturity T.*

Proof: By Theorem 3.5.1, $\lambda > 0$ and $\mu \geq 0$ for $\frac{x_0}{E[\rho(T)]} \leq z < \frac{x_0}{a}$. Thus the result follows immediately from Theorem 3.6.1 $\square$

The following lemma implies that the portion of the variance minimizing frontier corresponding to $\frac{x_0}{E[\rho(T)]} \leq z < \frac{x_0}{a}$ is *exactly* the efficient frontier that we are ultimately interested in.

Lemma 3.6.1 *Denote by $J_1^*(z)$ the optimal value of (3.6) corresponding to $z > 0$, where $a < \frac{x_0}{z} < b$. Then $J_1^*(z)$ is strictly increasing for $z \in [\frac{x_0}{E[\rho(T)]}, \frac{x_0}{a})$, and strictly decreasing for $z \in (\frac{x_0}{b}, \frac{x_0}{E[\rho(T)]}]$.*

Proof: For any z_1 and z_2 with $\frac{x_0}{a} > z_2 > z_1 \geq z_0 := \frac{x_0}{E[\rho(T)]}$, denote by $x_i^*(\cdot)$ the optimal wealth process of (3.6) corresponding to z_i, $i = 0, 1, 2$. Notice that z_1 can be represented as

$$z_1 = kz_2 + (1-k)z_0,$$

where $k := \frac{z_1 - z_0}{z_2 - z_0} \in [0, 1)$. Define

$$x(t) := kx_2^*(t) + (1-k)x_0^*(t), \quad \forall t \in [0, T].$$

Then $x(\cdot)$ is a feasible wealth process corresponding to z_1 due to the linearity of the system (2.5). Thus, noting that $0 \leq k < 1$,

$$J_1^*(z_1) \leq \text{Var } x(T) = k^2 \text{Var } x_2^*(T) < J_1^*(z_2).$$

This shows that $J_1^*(z)$ is strictly increasing for $z \in [\frac{x_0}{E[\rho(T)]}, \frac{x_0}{a})$. Similarly we can prove the second assertion of the lemma. $\square$

We are now ready to state the final result of this section.

Theorem 3.6.3 *Let x_0 be fixed. The efficient frontier for (3.36) is determined by the following parameterized equations:*

$$\begin{cases} E[x^*(T)] = z, \\ \text{Var } x^*(T) = \lambda(z)z - \mu(z)x_0 - z^2, \qquad \frac{x_0}{E[\rho(T)]} \le z < \frac{x_0}{a}, \end{cases} \tag{3.41}$$

where $(\lambda(z), \mu(z))$ is the unique solution to (3.18) (parameterized by z). Moreover, all the efficient portfolios are those variance minimizing portfolios corresponding to $z \in [\frac{x_0}{E[\rho(T)]}, \frac{x_0}{a})$.

Proof: First let us determine the variance minimizing frontier. Let $x^*(\cdot)$ be the wealth process under the variance minimizing portfolio corresponding to $z = E[x^*(T)]$. Then

$$\begin{aligned} \text{Var } x^*(T) = \; & E[x^*(T)^2] - z^2 \\ = \; & E[(\lambda(z) - \mu(z)\rho(T))x^*(T)] - z^2 \\ = \; & \lambda(z)E[x^*(T)] - \mu(z)E[\rho(T)x^*(T)] - z^2 \\ = \; & \lambda(z)z - \mu(z)x_0 - z^2, \end{aligned}$$

where the second equality followed from the general fact that $x^2 = \alpha x$ if $x = \alpha^+$. Now, Lemma 3.6.1 yields that the efficient frontier is the portion of the variance minimizing frontier corresponding to $\frac{x_0}{E[\rho(T)]} \le z < \frac{x_0}{a}$. This completes the proof. □

We remark that for z as in (3.41) the equality $Ex(T) = z$ in (3.6) can be replaced by the inequality $Ex(T) \ge z$, and one will get the same solution.

To conclude this section, we remark that the approaches and results of this chapter on the no-bankruptcy problem (3.6) can easily be adapted to the problem with a *benchmark floor*:

$$\begin{array}{ll} \text{Minimize} & \text{Var } x(T) \equiv Ex(T)^2 - z^2, \\ \text{subject to} & \begin{cases} Ex(T) = z, \\ x(t) \ge \underline{x}(t), \quad \text{a.s.}, \\ \pi(\cdot) \in L^2_{\mathcal{F}}(0,T;\mathbf{R}^m), \\ (x(\cdot), \pi(\cdot)) \text{ satisfies equation (2.5)}, \end{cases} \end{array} \tag{3.42}$$

where $\underline{x}(\cdot)$ is the wealth process of a *benchmark portfolio* (which is an admissible portfolio but not necessarily starting with the same initial wealth x_0).

For the model (3.42) the condition $x(t) \geq \underline{x}(t)$ implies that $\underline{x}(0) \leq x_0$ and $E\underline{x}(T) \leq z$. A similar argument as in the simplification of the no bankruptcy constraint yields that this condition is equivalent to $x(T) \geq \underline{x}(T)$.

The counterpart of problem (3.8) corresponding to problem (3.42) is

$$\begin{array}{ll} \text{Minimize} & EX^2 - z^2, \\ \text{subject to} & \begin{cases} EX = z, \\ E[\rho(T)X] = x_0, \\ X \in L^2(\mathcal{F}_T, \mathbf{R}), \quad X \geq \underline{x}(T), \text{ a.s..} \end{cases} \end{array} \tag{3.43}$$

The above problem is equivalent to

$$\begin{array}{ll} \text{Minimize} & E[Y + \underline{x}(T)]^2 - z^2, \\ \text{subject to} & \begin{cases} EY = \underline{z}, \\ E[\rho(T)Y] = \underline{y}_0, \\ Y \in L^2(\mathcal{F}_T, \mathbf{R}), \quad Y \geq 0, \text{ a.s.,} \end{cases} \end{array} \tag{3.44}$$

where $\underline{z} = z - E\underline{x}(T)$ and $\underline{y}_0 = x_0 - \underline{x}(0)$. Compared with problem (3.8), the cost function of (3.44) involves a first-order term of Y. However, (3.44) can be readily solved using exactly the same approach as in the proof of Theorem 3.4.1.

An interesting special case of this model is when $\underline{x}(T) = \underline{x}_T$, where $\underline{x}_T$ is a positive deterministic constant. In this case $\underline{x}(\cdot)$ is the wealth process under a risk-free portfolio (similar to the one in Theorem 3.6.2) with the terminal wealth $\underline{x}_T$ (alternatively, one may regard $\underline{x}(t) = \underline{x}_T B(t,T)$ where $B(t,T)$ is the time-t price of a unit discount Treasury bond maturing at time T). Thus, the process $\underline{x}(\cdot)$ provides a natural floor for the wealth process of an investor who wishes that his/her terminal wealth is at least $\underline{x}_T$ with probability one.

Obviously, the benchmark portfolio cannot be chosen arbitrarily. It must be selected so that the above problem is feasible. A feasibility study similar to the one in Section 3 will lead to proper conditions.

3.7 Special case of deterministic market coefficients

For the general case of a market with random coefficients, we have (see Proposition 3.6.1) derived the efficient portfolios as ones that replicate certain European put options with exercise price λ and expiration date T and written on a fictitious security having time-T price $\mu\rho(T)$. Moreover, to find this replicating portfolio it suffices to find a trading strategy $\pi^*(\cdot)$ along with a wealth process $x^*(\cdot)$ satisfying the BSDE

$$\begin{cases} dx^*(t) = [r(t)x^*(t) + \pi^*(t)'B(t)]dt + \pi^*(t)'\sigma(t)dW(t), \\ x^*(T) = (\lambda - \mu\rho(T))^+. \end{cases} \tag{3.45}$$

By the BSDE theory we know there exist a unique admissible portfolio $\pi^*(\cdot)$ along with a wealth process $x^*(\cdot)$ satisfying this BSDE, but actually solving this BSDE is sometimes easier said than done. This is because, in general, one is not able to express $(x^*(\cdot), \pi^*(\cdot))$ in a closed form. However, if all the market coefficients are deterministic (albeit time-varying), then, as will be shown in this section, an explicit form for $(x^*(\cdot), \pi^*(\cdot))$ is obtainable. In particular, we shall obtain analytical representations of the efficient portfolios via the Black–Scholes equation.

Throughout this section, in addition to all the basic assumptions specified earlier, we assume that $r(\cdot)$ and $\theta(\cdot)$ are deterministic functions (although $b(\cdot)$ and $\sigma(\cdot)$ themselves do not need to be deterministic). Notice that, according to Theorem 3.6.3 in the present case, the efficient portfolios are the variance minimizing portfolios corresponding to $z \geq x_0 e^{\int_0^T r(s)ds}$.

Theorem 3.7.1 *Assume that $\int_0^T |\theta(t)|^2 dt > 0$. Then there is a unique efficient portfolio for (3.6) corresponding to any given $z \geq x_0 e^{\int_0^T r(s)ds}$. Moreover, the efficient portfolio and the associated wealth process are given respectively as*

$$\begin{aligned} \pi^*(t) = &\ N(-d_+(t, y(t)))(\sigma(t)\sigma(t)')^{-1}B(t)y(t) \\ = &\ -(\sigma(t)\sigma(t)')^{-1}B(t)[x^*(t) - \lambda N(-d_-(t, y(t)))e^{-\int_t^T r(s)ds}] \end{aligned} \tag{3.46}$$

and

$$x^*(t) = \lambda N(-d_-(t, y(t)))e^{-\int_t^T r(s)ds} - N(-d_+(t, y(t)))y(t), \tag{3.47}$$

where $N(\cdot)$, with $N(x) := \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{v^2}{2}} dv$, is the cumulative distribution function of the standard normal distribution,

$$\begin{aligned} & y(t) := \mu \exp\{-\textstyle\int_0^T [2r(s) - |\theta(s)|^2]ds\} \exp\{\int_0^t [r(s) - \frac{3}{2}|\theta(s)|^2]ds - \int_0^t \theta'(s)dW(s)\}, \\ & d_+(t, y) := \frac{\ln(y/\lambda) + \int_t^T [r(s) + \frac{1}{2}|\theta(s)|^2]ds}{\sqrt{\int_t^T |\theta(s)|^2 ds}}, \\ & d_-(t, y) := d_+(t, y) - \sqrt{\textstyle\int_t^T |\theta(s)|^2 ds}, \end{aligned} \tag{3.48}$$

and (λ, μ), *with* $\lambda > 0, \mu \geq 0$, *is the unique solution to (3.18).*

Proof: First of all, in view of Remark 3.3.1, $a = 0$ and $b = +\infty$ under the given assumptions. Moreover, taking expectation on equation (3.2) and solving the resulting ordinary differential equation we get immediately that $E[\rho(T)] = e^{-\int_0^T r(s)ds}$. Thus a specialization of Theorem 3.6.3 establishes that the unique efficient portfolio exists for (3.6) corresponding to any $z \geq x_0 e^{\int_0^T r(s)ds}$.

Now consider the fictitious security process $y(\cdot)$ explicitly given in (3.48). Ito's formula shows that $y(\cdot)$ satisfies

$$\begin{cases} dy(t) = y(t)[(r(t) - |\theta(t)|^2)dt - \theta(t)'dW(t)], \\ y(0) = \mu \exp\{-\int_0^T [2r(s) - |\theta(s)|^2]ds\}, \quad y(T) = \mu\rho(T). \end{cases} \tag{3.49}$$

By virtue of Proposition 3.6.1, the efficient portfolio $\pi^*(\cdot)$ corresponding to a $z \geq x_0 e^{\int_0^T r(s)ds}$ is a replicating portfolio for a European put option written on $y(\cdot)$ with the strike λ and expiration date T. Now, we need to find $(x^*(\cdot), \pi^*(\cdot))$ that satisfies (3.45). Write $x^*(t) = f(t, y(t))$ for some function $f(\cdot, \cdot)$ (to be determined). Applying Ito's formula to f and (3.49) and then comparing with (3.45) in terms of both the drift and diffusion terms, we obtain

$$\pi^*(t) = -(\sigma(t)\sigma(t)')^{-1}B(t)\frac{\partial f}{\partial y}(t, y(t))y(t), \tag{3.50}$$

whereas f satisfies the following partial differential equation

$$\begin{cases} \frac{\partial f}{\partial t}(t, y) + r(t)y\frac{\partial f}{\partial y}(t, y) + \frac{1}{2}|\theta(t)|^2 y^2 \frac{\partial^2 f}{\partial y^2}(t, y) = r(t)f(t, y), \\ f(T, y) = (\lambda - y)^+. \end{cases} \tag{3.51}$$

This is exactly the Black–Scholes equation (generalized to deterministic but not necessarily constant coefficients) for a European put option; hence one can write down its solution explicitly as [4]

$$f(t, y) = \lambda N(-d_-(t, y))e^{-\int_t^T r(s)ds} - N(-d_+(t, y))y. \tag{3.52}$$

Finally, simple (yet non-trivial) calculations lead to

$$\frac{\partial f}{\partial y}(t, y) = -N(-d_+(t, y)).$$

[4] There are at least two ways to obtain the solution (3.52). One is to use the more familiar European call option formula and then use the put-call parity. The other is simply to check that the solution given by (3.52) indeed satisfies the Black–Scholes equation (3.51).

Thus the desired results (3.46) and (3.47) follow from (3.50) as well as the fact that $x^*(t) = f(t, y(t))$. □

Remark 3.7.1 The second expression of the efficient portfolio in (3.46) is in a feedback form, namely, it is a function of the wealth. In the case where bankruptcy is allowed (see Zhou and Li [70]), the efficient portfolio is

$$\pi^*(t) = -(\sigma(t)\sigma(t)')^{-1}B(t)[x^*(t) - \gamma e^{-\int_t^T r(s)ds}], \tag{3.53}$$

where

$$\gamma := \frac{z - x_0 e^{\int_0^T [r(t) - |\theta(t)|^2]dt}}{1 - e^{-\int_0^T |\theta(t)|^2 dt}}$$

Note the striking resemblance in form between (3.46) and (3.53).

Remark 3.7.2 The discounted price process of any financial security must be a martingale under the risk neutral probability measure Q. Since it can be easily verified that the process $y(\cdot)$ given in (3.48) satisfies $y(T) = \mu\rho(T)$ and $y(t) = S_0(t)E_Q[S_0(T)^{-1}y(T)|\mathcal{F}_t]$ for $t \in [0, T]$, it follows that the process $y(\cdot)$ can be interpreted as the price process of a *fictitious* security that takes the value $\mu\rho(T)$ at the maturity date T. We say fictitious security, as the price process $y(\cdot)$ does not belong to our underlying market, which is comprised of the securities with price processes $S_i(\cdot),\ i = 0, 1, 2, \ldots, m$.

Remark 3.7.3 It appears that expression (3.46) for the optimal trading strategy $\pi^*(\cdot)$ is not convenient for practical implementation because it is in terms of the fictitious security process $y(\cdot)$ which, in fact, is not directly observable. There are at least two ways to deal with this issue. First, simple manipulation shows that equation (3.49) is nothing else but the wealth equation (2.5) under the portfolio

$$\hat{\pi}(t) := \quad -(\sigma(t)\sigma(t)')^{-1}B(t)y(t). \tag{3.54}$$

Notice that, with the initial endowment $y(0) = \mu \exp\{-\int_0^T [2r(s) - |\theta(s)|^2]ds\}$, the above $\hat{\pi}(\cdot)$ is a legitimate, implementable continuous-time portfolio because it is a feedback of the wealth process $y(\cdot)$. The portfolio $\hat{\pi}(\cdot)$ is also called a (continuous-time) *mutual fund* or a *basket of stocks*. Thus, one may compose, actually or virtually (via a simulation, say), a portfolio using the initial wealth $y(0)$ and the strategy $\hat{\pi}(\cdot)$, and the corresponding wealth process as determined via (2.5) is exactly the fictitious security process $y(\cdot)$ which is observable. The efficient portfolio is then the replicating portfolio for a European put option (with strike λ and maturity T) written on this basket of stocks. Another way is based on the observation that, since the market is complete, the "auxiliary" process $y(\cdot)$ can be inferred from the returns of the risky securities. To see this, define $DS(t) := (\frac{dS_1(t)}{S_1(t)}, \ldots, \frac{dS_m(t)}{S_m(t)})'$ and $b(t) := (b_1(t), \ldots, b_m(t))'$. Then one can solve for $dW(t)$ from equation (2.2), obtaining

$$dW(t) = \sigma(t)^{-1}\big[DS(t) - b(t)dt\big].$$

Consequently, one can compute the value of $y(t)$ for every $t \geq 0$ by combining the above with (3.48). In practice, this can provide an approximation of $y(\cdot)$ in terms of discrete–time asset returns.

Remark 3.7.4 Continuing with the second approach discussed in the preceding remark, we can express the fictitious process $y(\cdot)$ explicitly in terms of the stock prices if all the coefficients are time-invariant. In fact, in this case Ito's formula yields

$$\begin{aligned}
\ln S_i(t) - \ln S_i(0) &= (b_i - \frac{1}{2}\sum_{j=1}^{m}|\sigma_{ij}|^2)t + \sum_{j=1}^{m}\sigma_{ij}W^j(t) \\
&= (r - \frac{1}{2}\sum_{j=1}^{m}|\sigma_{ij}|^2)t + (b_i - r)t + \sum_{j=1}^{m}\sigma_{ij}W^j(t).
\end{aligned}$$

Solving for $W(t)$ we get

$$W(t) = \sigma^{-1}V(t) - \theta' t$$

where $V(t) := (v_1(t), \cdots, v_m(t))'$ with $v_i(t) := \ln S_i(t) - \ln S_i(0) - (r - \frac{1}{2}\sum_{j=1}^{m}|\sigma_{ij}|^2)t$. Substituting the above to (3.48) we obtain

$$\begin{aligned}
y(t) &= y(0)\exp\{(r - \frac{3}{2}|\theta|^2)t - \theta W(t)\} \\
&= y(0)\exp\{(r - \frac{1}{2}|\theta|^2)t - \theta\sigma^{-1}V(t)\}.
\end{aligned}$$

In particular, in the simple Black–Scholes case where the interest rate is constant and there is a single risky asset whose price process $S_1(\cdot)$ is taken as geometric Brownian motion: $S_1(t) = S_1(0)\exp\{(b - \sigma^2/2)t + \sigma W(t)\}$, by the preceding formula the fictitious security process is of the form $y(t) = \alpha e^{\beta t}[S_1(t)]^{-\theta/\sigma}$, where $\alpha > 0$ and β are two computable scalars. But since $\theta > 0$ and $\sigma > 0$ it is apparent that this contingent claim has a positive payoff (i.e., is "in the money") if and only if the terminal price $S_1(T)$ is greater than some positive constant (the "strike price"). In this respect the contingent claim resembles a conventional call, and it is in accordance with economic intuition: the bigger the terminal price $S_1(T)$ of the risky asset, the better for the investor.

Remark 3.7.5 The terminal wealth under an efficient portfolio is of the form $(\lambda - \mu\rho(T))^+$, which may take zero value with *positive* probability. Nevertheless, by risk neutral valuation for each $t < T$ the portfolio value is strictly positive with probability one, and so a trading strategy that replicates this contingent claim is well-defined for $t < T$ as a proportional strategy. However, for the reasons discussed in Section 3.2, it is not clear whether such a proportional strategy will satisfy a reasonable condition of admissibility, such as the ones found in Cvitanic and Karatzas [10] and Karatzas and Shreve [26].

In Theorem 3.7.1, (λ, μ) is the unique solution to (3.18), a solution that is ensured by Theorem 5.1. It turns out that, in the case of deterministic coefficients, (3.18) has a more explicit form.

Proposition 3.7.1 *Under the assumptions of Theorem 3.7.1, if* $z > x_0 e^{\int_0^T r(s)ds}$*, then* (λ, μ) *is the unique solution to the following system of equations:*

$$\begin{cases} \lambda N\left(\frac{\ln\frac{\lambda}{\mu}+\int_0^T[r(s)-\frac{1}{2}|\theta(s)|^2]ds}{\sqrt{\int_0^T|\theta(s)|^2ds}}\right) - \mu e^{\int_0^T[|\theta(s)|^2-r(s)]ds} N\left(\frac{\ln\frac{\lambda}{\mu}+\int_0^T[r(s)-\frac{3}{2}|\theta(s)|^2]ds}{\sqrt{\int_0^T|\theta(s)|^2ds}}\right) = x_0 e^{\int_0^T r(s)ds}, \\ \lambda N\left(\frac{\ln\frac{\lambda}{\mu}+\int_0^T[r(s)+\frac{1}{2}|\theta(s)|^2]ds}{\sqrt{\int_0^T|\theta(s)|^2ds}}\right) - \mu e^{-\int_0^T r(s)ds} N\left(\frac{\ln\frac{\lambda}{\mu}+\int_0^T[r(s)-\frac{1}{2}|\theta(s)|^2]ds}{\sqrt{\int_0^T|\theta(s)|^2ds}}\right) = z. \end{cases} \tag{3.55}$$

Proof: First note that when $z > x_0 e^{\int_0^T r(s)ds}$, it follows from Theorem 3.5.1 that $\lambda > 0$ and $\mu > 0$. We start with the second equation in (3.18):

$$E[\rho(T)(\lambda - \mu\rho(T))^+] = x_0. \tag{3.56}$$

By the proof of Theorem 3.7.1, $x_0 = x^*(0) = f(0, y(0))$. Using the expressions for $f(\cdot,\cdot)$ and $y(0)$ as given in (3.52) and (3.49) respectively, we conclude that $f(0, y(0))e^{\int_0^T r(s)ds}$ equals the left hand side of the first equation in (3.55). Hence the first equation in (3.55) follows.

Next, the first equation in (3.18) can be rewritten as

$$E[\rho(T)(\frac{\lambda}{\rho(T)} - \mu)^+] = z. \tag{3.57}$$

Drawing an analog between (3.57) and (3.56), we see that equation (3.57) is nothing else than a statement that z is the initial price of a European *call* option on $\frac{\lambda}{\rho(T)}$ with strike μ and maturity T. Define

$$\bar{y}(t) := \lambda \exp\{\int_0^t [r(s) + \frac{1}{2}|\theta(s)|^2]ds + \int_0^t \theta(s)' dW(s)\}, \tag{3.58}$$

which satisfies

$$\begin{cases} d\bar{y}(t) = \bar{y}(t)[(r(t) + |\theta(t)|^2)dt + \theta(t)dW(t)], \\ \bar{y}(0) = \lambda, \ \ \bar{y}(T) = \frac{\lambda}{\rho(T)}. \end{cases} \tag{3.59}$$

The well-known Black–Scholes call option formula (or going through a similar derivation to that in the proof of Theorem 3.7.1) implies that $z = g(0, \bar{y}(0))$ where

$$g(t, y) = N(\bar{d}_+(t, y))y - \mu N(\bar{d}_-(t, y))e^{-\int_t^T r(s)ds}, \tag{3.60}$$

with

$$\begin{aligned} \bar{d}_+(t, y) &:= \frac{\ln(y/\mu) + \int_t^T [r(s) + \frac{1}{2}|\theta(s)|^2]ds}{\sqrt{\int_t^T |\theta(s)|^2 ds}}, \\ \bar{d}_-(t, y) &:= \bar{d}_+(t, y) - \sqrt{\int_t^T |\theta(s)|^2 ds}. \end{aligned} \tag{3.61}$$

This leads to the second equation in (3.55). $\square$

We now turn to the representation of the efficient frontier. For the general case this is provided by Theorem 3.6.3, where we represented the minimal variance Var $x^*(T)$ as a function of the expected terminal wealth $E[x^*(T)]$ $(= z)$. But there is the drawback to representation (3.41) in Theorem 3.6.3, namely, the minimal variance Var $x^*(T)$ is an implicit function of z, because the Lagrange multipliers $\lambda(z)$ and $\mu(z)$ are, in general, implicit functions of z. It turns out that in the deterministic coefficient case we can write the efficient frontier in an explicit parametric form, as a function of a positive scalar variable that we denote by η.

Theorem 3.7.2 *Under the assumptions of Theorem 3.7.1, the efficient frontier is the* *following*

$$\left\{\begin{array}{rcl} E[x^*(T)] & = & \frac{\eta e^{\int_0^T r(t)dt} N_1(\eta)-N_2(\eta)}{\eta N_2(\eta)-e^{-\int_0^T [r(s)-|\theta(s)|^2]ds} N_3(\eta)} x_0, \\ \text{Var } x^*(T) & = & \left[\frac{\eta}{\eta N_1(\eta)-e^{-\int_0^T r(t)dt} N_2(\eta)} - 1\right] [Ex^*(T)]^2 \\ & & -\frac{x_0}{\eta N_1(\eta)-e^{-\int_0^T r(t)dt} N_2(\eta)} Ex^*(T), \quad \eta \in (0,\infty], \end{array}\right. \tag{3.62}$$

where

$$\begin{array}{l} N_1(\eta) := N\left(\frac{\ln\eta+\int_0^T [r(s)+\frac{1}{2}|\theta(s)|^2]ds}{\sqrt{\int_0^T |\theta(s)|^2 ds}}\right), \\ N_2(\eta) := N\left(\frac{\ln\eta+\int_0^T [r(s)-\frac{1}{2}|\theta(s)|^2]ds}{\sqrt{\int_0^T |\theta(s)|^2 ds}}\right), \\ N_3(\eta) := N\left(\frac{\ln\eta+\int_0^T [r(s)-\frac{3}{2}|\theta(s)|^2]ds}{\sqrt{\int_0^T |\theta(s)|^2 ds}}\right). \end{array} \tag{3.63}$$

Proof: Set $\eta := \frac{\lambda}{\mu}$. From the proof of Theorem 3.5.1, it follows that as z runs from $x_0 e^{\int_0^T r(t)dt}$ (inclusive) to ∞ (exclusive), η changes decreasingly from ∞ (inclusive) to 0 (exclusive). Therefore $\eta \in (0,\infty]$. Dividing the second equation by the first one in (3.55) we get the first equation of (3.62). Now, replacing λ by $\eta\mu$ in the second equation of (3.55) and solving for μ, we obtain

$$\mu = \frac{z}{\eta N_1(\eta) - e^{-\int_0^T r(t)dt} N_2(\eta)}. \tag{3.64}$$

Thus, appealing to (3.41), we have

$$\text{Var } x^*(T) = \lambda z - \mu x_0 - z^2 = \eta\mu z - z^2 - \mu x_0.$$

Using (3.64) and noting $z \equiv E[x^*(T)]$, we get the second equation of (3.62). □

Remark 3.7.6 Although the efficient frontier does not have a closed analytical form, equation (3.62) is "explicit" enough in the sense that it has only one parameter $\eta \in (0,\infty]$. It is easy to numerically draw the curve based on (3.62).

Analogous to the single-period case, the efficient frontier in continuous time will induce the so-called *capital market line* (CML). Specifically, define $r^*(t) := \frac{x^*(t)-x_0}{x_0}$, the return rate of an efficient strategy at time t. Then in the case where bankruptcy is allowed, the capital market line is the following straight line in the terminal mean–standard deviation plane (see Zhou [69]):

$$Er^*(T) = r_f(T) + \sqrt{e^{\int_0^T |\theta(t)|^2 dt} - 1}\, \sigma_{r^*(T)}, \tag{3.65}$$

where $r_f(T) := e^{\int_0^T r(t)dt} - 1$ is the risk-free return rate over $[0,T]$, and $\sigma_{r^*(T)}$ denotes the standard deviation of $r^*(T)$. In the present case of bankruptcy prohibition, we can easily obtain the corresponding CML via the efficient frontier (3.62). Clearly the CML is no longer a straight line, as seen from (3.62).

Example 3.7.1 Take the same example as in Zhou and Li [70] where a market has a bank account with $r(t) = 0.06$ and only one stock with $b(t) = 0.12$ and $\sigma(t) = 0.15$. An agent starts with an endowment $x_0 = \$1$ million and expects a terminal mean payoff $z = \$1.2$ million at $T = 1$ (year). Bankruptcy is not allowed (as opposed to Zhou and Li [70]). In this case $\theta(t) = 0.4$. Thus the system of equations (3.55) reduces to

$$\begin{cases} \lambda N\left(\frac{\ln(\lambda/\mu)-0.02}{0.4}\right) - \mu e^{0.1} N\left(\frac{\ln(\lambda/\mu)-0.18}{0.4}\right) = e^{0.06}, \\ \lambda N\left(\frac{\ln(\lambda/\mu)+0.14}{0.4}\right) - \mu e^{-0.06} N\left(\frac{\ln(\lambda/\mu)-0.02}{0.4}\right) = 1.2. \end{cases} \tag{3.66}$$

Solving this equation we get

$$\lambda = 2.0220, \quad \mu = 0.8752.$$

Therefore the corresponding efficient portfolio is the replicating portfolio of a European put option on the following fictitious stock

$$\begin{cases} dy(t) = y(t)[-0.1dt - 0.4dW(t)], \\ y(0) = \$0.9109 \end{cases} \tag{3.67}$$

with a strike price \$2.0220 maturing at the end of the year.

The CML when bankruptcy is allowed has been obtained in Zhou and Li (2000) as

$$Er^*(1) = 0.0618 + 0.4165\sigma_{r^*(1)}. \tag{3.68}$$

In the current case of no bankruptcy, the CML is the following based on (3.62):

$$\begin{aligned} Er^*(1) &= \frac{e^{0.06}\eta N(\frac{\ln\eta+0.14}{0.4}) - N(\frac{\ln\eta-0.02}{0.4})}{\eta N(\frac{\ln\eta-0.02}{0.4}) - e^{0.1}N(\frac{\ln\eta-0.18}{0.4})} - 1, \\ \sigma^2_{r^*(1)} &= [\frac{\eta}{\eta N(\frac{\ln\eta+0.14}{0.4}) - e^{-0.06}N(\frac{\ln\eta-0.02}{0.4})} - 1][Er^*(1)+1]^2 \\ &\quad - \frac{1}{\eta N(\frac{\ln\eta+0.14}{0.4}) - e^{-0.06}N(\frac{\ln\eta-0.02}{0.4})}[Er^*(1)+1]. \end{aligned} \tag{3.69}$$

Both (3.68) and (3.69) are plotted on the same plane; see Figure 3.1. We see that (3.69) falls below (3.68), which is certainly expected. In particular, if an agent is expecting an annual return rate of 20%, then the corresponding standard deviation with bankruptcy allowed is 33.1813%, whereas the one without bankruptcy is 33.3540%.

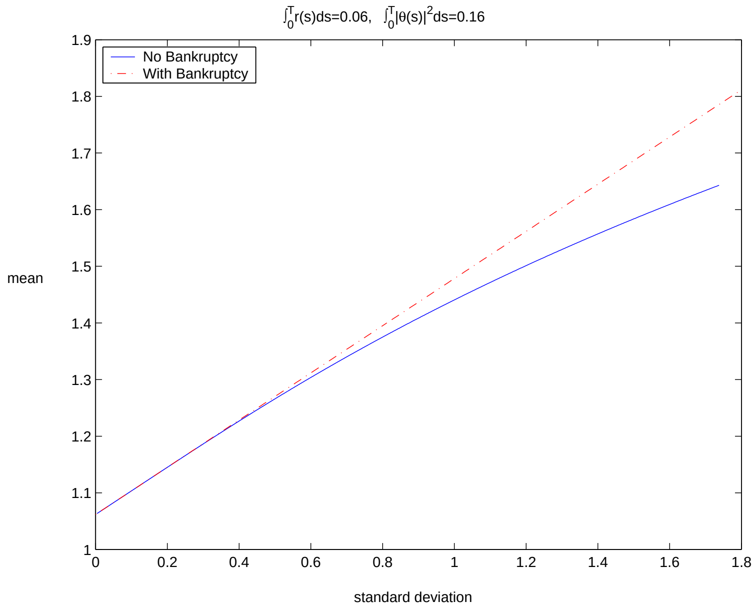

Figure 3.1: Mean-standard deviation of the terminal return rate (CML)

3.8 Concluding remarks

In a complete financial market, the continuous mean-variance problem with unconstrained portfolios was solved out completely by Zhou and Li [70]. In this chapter, we investigate a continuous-time mean–variance portfolio selection problem with stochastic parameters under a no bankruptcy constraint. The problem has been completely solved. The main idea is the decomposition of the continuous-time portfolio selection problem. We first identify the optimal terminal wealth attainable by those constrained portfolios, and then replicate this optimal wealth. This idea in fact applies to a more general class of constrained continuous-time portfolio selection problem: first translate all the constraints to the ones imposed on the terminal wealth, solve this constrained optimization problem on random variables, and then replicate the contingent claim represented by the optimal terminal wealth. As we showed in Section 3.4, we can easily apply this idea to solve the unconstrained mean-variance problem. Later on, we will apply this idea for incomplete market and other portfolio optimization problem as well.

As we emphasize in Section 3.2 and elsewhere, by defining trading strategies in terms of the amount of money invested in individual assets, rather than in terms of the proportion of wealth invested in individual assets, we can allow for strategies where the portfolio's value becomes zero before the terminal date with positive probability. Hence our approach, which includes an explicit constraint on nonnegative portfolio value, leads to a strictly bigger set of admissible trading strategies than with the proportional strategy approach. It is an open question whether this larger class of admissible strategies gives a strictly better value of the optimal objective value than with the smaller class, although we conjecture that the two values are the same. And if the two optimal objective values are indeed the same, it is another open question whether this common value is attained by some proportional trading strategy. This is an open question because if you try to convert our optimal strategy to a proportional strategy, then it might be well defined for $t < T$, but even so it might not be admissible because the ratio of the money in a risky asset to the total wealth might not be well-behaved. Since the optimal attainable wealth takes the value zero with positive probability, it is clear it cannot be replicated by a proportional trading strategy satisfying the admissibility condition given immediately before (2.8). However, since the attainable wealth process is strictly positive with probability one for all $t < T$, it is an open question whether some other reasonable definition of admissibility might lead to a proportional trading strategy that does replicate the optimal attainable wealth.

In some financial market there is no-shorting constraint on risky securities. This is another type of constraint. When shorting is prohibited, there are a lot of un-replicable contingent claims, therefore the market is essentially incomplete. We will study this constraint, along with the constraints combined by shorting prohibition and bankruptcy prohibition, in the next chapter.

Chapter 4

Mean-Variance Criteria in an Incomplete Market

4.1 Introduction

In Chapter 3, we investigated the mean-variance problem in a complete market. In that chapter, the completeness of the market is essential, as there is a unique equivalent martingale measure in the market, and therefore one can completely transfer the complex dynamic portfolio selection problem into a simpler static optimization problem.

Completeness of the market is an ideal assumption for the financial market. But to be more practical, we cannot always suppose the market is complete. A fundamental fact is that the number of the securities may not be the same as that of the random resources (the dimension of the Brownian motion). Even if there are sufficient securities being traded in the market, it is not realistic to require the investors to invest in all the available securities. They may only concern on a subset of the securities in the market. Therefore, for these investors, the "market" is essentially incomplete. Another possible reason for the incompleteness of the market is the constraint, such as shorting prohibition, on the portfolio.

There are two ways to deal with portfolio selection problems in a continuous-time market. One is the "backward" method, like the one used in Chapter 3, which applies the martingale method to separate the problem into a static optimization problem and a backward stochastic differential equation. This method was developed by Pliska [53], Cox and Huang [8, 9], and Karatzas, Lehoczky and Shreve [24]. The other is the "forward" method, which applies the stochastic optimal control technique to the problem. This method was adopted by Li and Zhou [70], Hu and Zhou [20], Lim and Zhou [36], Li, Zhou and Lim [34], and Lim [35] in recent years for solving mean-variance problems. But when the market is incomplete, both approach require more complicated analysis. Karatzas, Lehoczky, Shreve and Xu [25] deal with the incomplete market by constructing some fictitious securities in the market and regarding the original market as a constrained market where the fictitious securities are prohibited to be traded, and then apply the martingale method to maximize the utility. The same technique was applied in Lim [35], where stochastic optimal control method was adopted. In Hu and Zhou [20], a type of so-called cone constraint was studied in a complete market, the result of which may also be applied to deal with the incomplete market.

Portfolio selection problem in continuous-time, especially in an incomplete market, has been studied by utility maximization framework in literature. In that mainly in the framework, the utility function is typically assumed to be strictly increasing, infinite derivative at 0, among others. These assumptions are so strict that even the quadratic function does not satisfy them. See Karatzas, Lehoczky, Shreve and Xu [25] and Schachermayer [60].

In this chapter, we will study the mean-variance portfolio selection problem in an incomplete market. The rest of this chapter is organized as follows: The incomplete market and the (constrained) mean-variance problem will be specified in Section 4.2. In Section 4.3, we investigate the unconstrained mean-variance problem, which is the simplest case in this chapter. We set up the equivalent static optimization problem and its dual problem. When the parameters are deterministic, we can solve the dual problem, and obtain the optimal solution for the mean-variance problem explicitly. In Section 4.4, we deal with the shorting prohibition model. Section 4.5 is devoted to no bankruptcy case. At last, we discuss the case where both shorting and bankruptcy are prohibited in Section 4.6. Although in the last 3 cases, the problem become very complicated, we work out the explicit optimal solutions when the parameters are deterministic. Finally we conclude the chapter in Section 4.7.

4.2 Problem formulation

Again we adopt the market given in Chapter 2. Contrary to Chapter 3, the market is not necessarily complete, and the number of risky securities n may be less than m, the dimension of the Brownian motion.

In this chapter, we make the following assumption to ensure that the market is arbitrage-free (see Theorem 2.3.2).

Assumption 4.2.1 *There exists* $\theta \in L^{\infty}_{\mathcal{F}}(0, T, \mathbf{R}^n)$ *such that* $\sigma(t)\theta(t) = B(t)$, *a.s., a.e.* $\in [0, T]$.

Assumption 4.2.1 is satisfied if $\sigma(t)'\sigma(t)$ is uniformly positive definite (i.e., there is $\delta > 0$ such that $\sigma(t)'\sigma(t) \geq \delta I_n$, *a.s.*, *a.e.t* $\in [0, T]$), in which case $n \leq m$ and there is a unique such θ. In general, however, the process θ, if it exists, may not be unique.

Define

$$\Theta := \{\theta \in L^{\infty}_{\mathcal{F}}(0, T, \mathbf{R}^n) : \sigma(t)\theta(t) = B(t),\ a.s.,\ a.e.t \in [0, T]\}, \tag{4.1}$$

and

$$\theta_0(t) := \text{argmin}_{\sigma(t)x=B(t)} |x|^2. \tag{4.2}$$

By Lemma A.2, θ_0 is a $\mathcal{F}_t$-progressively measurable process. Moreover, due to Assumption 4.2.1 we also have $\theta_0 \in L^{\infty}_{\mathcal{F}}(0, T, \mathbf{R}^n)$.

For any $\theta \in L^{\infty}_{\mathcal{F}}(0, T, \mathbf{R}^n)$, define

$$H_\theta(t) := \exp\left\{-\int_0^t [r(s) + \frac{1}{2}|\theta(s)|^2]ds - \int_0^t \theta(s)' dW(s)\right\}. \tag{4.3}$$

Equivalently, $H_\theta(\cdot)$ can be defined as the unique solution to the following SDE

$$\begin{cases} dH_\theta(t) = -r(t)H_\theta(t)dt - H_\theta(t)\theta(t)'dW(t), \\ H_\theta(0) = 1. \end{cases} \tag{4.4}$$

The following technical lemma is useful in the sequel.

Lemma 4.2.1 *Given a set* $A \subseteq L^{\infty}_{\mathcal{F}}(0, T, \mathbf{R}^n)$. *If* $k\theta_1 + (1-k)\theta_2 \in A$ *whenever* $\theta_1 \in A, \theta_2 \in A$ *and* $k \in L^1_{\mathcal{F}}(0, T; [0, 1])$, *then the set* $\{H_\theta(\cdot) : \theta \in A\}$ *is convex.*

Proof: For any $\theta_1, \theta_2 \in A$ and $\lambda \in [0,1]$, denote $H(\cdot) := \lambda H_{\theta_1}(\cdot) + (1-\lambda) H_{\theta_2}(\cdot)$. Then $H(0) = 1$, and

$$\begin{aligned} & dH(t) \\ = \;& \lambda[-r(t)H_{\theta_1}(t)dt - H_{\theta_1}(t)\theta_1(t)'dW(t)] + (1-\lambda)[-r(t)H_{\theta_2}(t)dt - H_{\theta_2}(t)\theta_2(t)'dW(t)] \\ = \;& -r(t)H(t)dt - H(t)[k(t)\theta_1(t) + (1-k(t))\theta_2(t)]'dW(t), \end{aligned}$$

where $k(t) := \frac{\lambda H_{\theta_1}(t)}{\lambda H_{\theta_1}(t) + (1-\lambda)H_{\theta_2}(t)}$. Define $\theta(t) := k(t)\theta_1(t) + [1-k(t)]\theta_2(t)$. Then $\theta \in A$. It then follows from the definition of H_θ, see (4.4), that $H(t) \equiv H_\theta(t)$. This completes the proof. □

Fix an initial wealth x_0. A general (constrained) continuous-time Markowitz's mean–variance portfolio selection problem is formulated as

$$\begin{array}{ll} \text{Minimize} & \text{Var}\, x(T) \equiv Ex(T)^2 - z^2, \\ \text{subject to} & \begin{cases} Ex(T) = z, \\ \pi(\cdot) \in \Pi, \\ (x(\cdot), \pi(\cdot)) \text{ satisfies equation (2.5) with } x(0) = x_0, \\ (x(\cdot), \pi(\cdot)) \in C, \end{cases} \end{array} \tag{4.5}$$

where C is a given convex set in $L^2_{\mathcal{F}}(0,T,\mathbf{R}) \times \Pi$, and $z \in \mathbf{R}$ is a parameter. The optimal portfolio for this problem (corresponding to a fixed z) is called an *efficient portfolio*, and the set of all points $(\text{Var}\; x^*(T), z)$, where $\text{Var}\; x^*(T)$ denotes the optimal value of (4.5) corresponding to z and z runs over $\mathbf{R}$, is called the *efficient frontier*.

In this chapter, the following four cases of the constraint set C will be studied respectively. It is easy to check that C is convex in each of the four cases.

Case 1. $C = L^2_{\mathcal{F}}(0,T,\mathbf{R}) \times \Pi$, corresponding to the case where portfolios are not constrained.

Case 2. $C = \{(x(\cdot), \pi(\cdot)) \in L^2_{\mathcal{F}}(0,T,\mathbf{R}) \times \Pi : \pi(t) \geq 0, \text{ a.s., a.e.} t \in [0,T]\}$, corresponding to the case where short-selling is prohibited.

Case 3. $C = \{(x(\cdot), \pi(\cdot)) \in L^2_{\mathcal{F}}(0,T,\mathbf{R}) \times \Pi : x(t) \geq 0, \text{ a.s., } \forall t \in [0,T]\}$, corresponding to the case where bankruptcy is prohibited.

Case 4. $C = \{(x(\cdot), \pi(\cdot)) \in L^2_{\mathcal{F}}(0,T,\mathbf{R}) \times \Pi : x(t) \geq 0, \text{ a.s., } \forall t \in [0,T]; \; \pi(t) \geq 0, \text{ a.s., a.e.} t \in [0,T]\}$, corresponding to the case where both bankruptcy and short-selling are prohibited.

Given a constraint set C assosicated with one of the four cases above, define the following *attainable terminal wealth set*:

$$
\begin{aligned}
A_C := \quad & \{X \in L^2(\mathcal{F}_T, \mathbf{R}) : \text{ there exist } x \in \mathbf{R} \text{ and } \pi(\cdot) \in \Pi \text{ such that } (x^\pi(\cdot), \pi(\cdot)) \\
& \text{satisfies (2.5) with } x^\pi(0) = x, \; x^\pi(T) = X, \text{ and } (x^\pi(\cdot), \pi(\cdot)) \in C\}.
\end{aligned}
\tag{4.6}
$$

Proposition 4.2.1 *A_C is a convex set in $L^2(\mathcal{F}_T, \mathbf{R})$.*

Proof: This is evident by virtue of the linearity of the wealth equation (2.5) and the convexity of the set C. □

By the decomposition approach, the following *static* optimization problem plays a critical role for solving problem (4.5):

$$
\begin{array}{ll}
\text{Minimize} & EX^2 - z^2, \\
\text{subject to} & \begin{cases} EX = z, \\ E[XH_{\theta_0}(T)] = x_0, \\ X \in A_C. \end{cases}
\end{array}
\tag{4.7}
$$

This problem is to locate the optimal attainable terminal wealth X^* in A_C. Once this is solved, an optimal portfolio for (4.5) can be obtained by replicating X^* (which is possible by the very definition of A_C). In Chapter 3, we found that when the market is complete, the static problem (4.7) is easier to solve than the original one because $A_C = L^2(\mathcal{F}_T, \mathbf{R})$. When the market is incomplete, the main difficulty is to characterize the attainable set A_C for each of the four cases.

The following result verifies that in order to solve the original problem (4.5) it suffices to solve (4.7).

Theorem 4.2.1 *If $(x^*(\cdot), \pi^*(\cdot))$ is optimal for (4.5), then $x^*(T)$ is optimal for (4.7). Conversely, if $X^* \in A_C$ is optimal for (4.7), then any wealth–portfolio pair $(x^*(\cdot), \pi^*(\cdot))$ satisfying (2.5), $(x^*(\cdot), \pi^*(\cdot)) \in C$ and $x^*(T) = X^*$ is optimal for (4.5).*

Proof: . This is straightforward by the definition of A_C. $\square$

Thanks to the convexity of the attainable set A_C, we can transform problem (4.7) by Proposition 3.4.1 to an equivalent problem as stipulated in the following theorem.

Theorem 4.2.2 *If problem (4.7) admits a solution X^*, then there exists a pair of scalars (λ, μ) such that X^* is also the optimal solution for the following problem:*

$$\begin{array}{ll} \text{Minimize} & E[X - (\lambda - \mu H_{\theta_0}(T))]^2, \\ \text{subject to} & X \in A_C. \end{array} \tag{4.8}$$

Conversely, if there is a pair of scales (λ, μ), such that the optimal solution X^* of (4.8) satisfies

$$\begin{cases} EX^* = z, \\ E[X^* H_{\theta_0}(T)] = x_0. \end{cases} \tag{4.9}$$

then X^* must be an optimal solution of (4.7).

This theorem suggests that one can first solve problem (4.8) for general (λ, μ), and then determine the values of (λ, μ) via the equations (4.9).

In the next four sections, we will study the four cases respectively. We will mainly devote ourselves to characterizing the attainable set A_C and solving (4.8) for each case. For the general situation when the market parameters $r(\cdot)$, $\mu_i(\cdot)$ and $\sigma_{ij}(\cdot)$ are stochastic processes, it is very difficult to solve (4.8) *explicitly* in terms of (λ, μ). However, for the case when all the market parameters are deterministic, we will obtain analytical solution to (4.8) and thereby get explicit solution to the original problem (4.5) for all the four cases.

4.3 Case 1: Unconstrained case

In this case the constraint set $C = L^2_{\mathcal{F}}(0, T, \mathbf{R}) \times \Pi$. Our first result characterizes the attainable terminal wealth set A_C for this constraint set.

Theorem 4.3.1 *For any $X \in L^2(\mathcal{F}_T, \mathbf{R})$, $X \in A_C$ if and only if $E[XH_\theta(T)]$ is independent of $\theta \in \Theta$.*

Proof: If $X \in A_C$, then there is $x_0 \in \mathbf{R}$ and a portfolio $\pi(\cdot) \in \Pi$ such that

$$\begin{cases} dx(t) = [r(t)x(t) + B(t)'\pi(t)]dt + \pi(t)'\sigma(t)dW(t), \\ x(0) = x_0, \;\; x(T) = X. \end{cases}$$

Now, for any $\theta \in \Theta$,

$$\begin{aligned} dx(t) &= [r(t)x(t) + B(t)'\pi(t)]dt + \pi(t)'\sigma(t)dW(t) \\ &= [r(t)x(t) + \theta(t)'\sigma(t)'\pi(t)]dt + \pi(t)'\sigma(t)dW(t). \end{aligned}$$

Applying Ito's formula, we obtain

$$x_0 \equiv x(0) = E[x(T)H_\theta(T)] = E[XH_\theta(T)],$$

implying that $E[XH_\theta(T)]$ is independent of the choice of $\theta \in \Theta$.

Conversely, assume that $E[XH_\theta(T)]$ does not depend on $\theta \in \Theta$. By the BSDE theory, for any $\theta \in \Theta$, the following equation

$$\begin{cases} dX(t) = [r(t)X(t) + \theta(t)'Z(t)]dt + Z(t)'dW(t), \\ X(T) = X \end{cases} \tag{4.10}$$

admits a unique solution pair $(X_\theta(\cdot), Z_\theta(\cdot))$, with $X_\theta(0) = E[XH_\theta(T)]$. So by the assumption $X_\theta(0)$, $\theta \in \Theta$, are all the same, which is denoted by x_0.

Next, fix $\theta \in \Theta$ and let $(X_\theta(\cdot), Z_\theta(\cdot))$ solves (4.10). We are to prove that there exist a portfolio $\pi_0(\cdot) \in \Pi$ such that

$$Z_\theta(t) = \sigma(t)'\pi_0(t), \text{ a.s., a.e.} t \in [0, T]. \tag{4.11}$$

Indeed, define

$$\pi_0(t) := \mathrm{argmin}_{\pi \in \mathrm{argmin}_{\pi \in \mathbf{R}^m} |\sigma(t)'\pi - Z_\theta(t)|^2} |\pi|^2.$$

Appealing to Lemma B.2 and Lemma A.2, $\pi_0(\cdot)$ is well-defined and it is a progressively measurable stochastic processes with respect to $\mathcal{F}_t$.

Set $\bar{D}(t) := \sigma(t)'\pi_0(t) - Z_\theta(t)$ and

$$D(t) := \begin{cases} 0, & \text{if } \bar{D}(t) = 0, \\ \bar{D}(t)/|\bar{D}(t)|, & \text{if } \bar{D}(t) \neq 0. \end{cases}$$

Then $D(\cdot) \in L^\infty_{\mathcal{F}}(0, T, \mathbf{R}^n)$. Moreover, $\sigma(t)\bar{D}(t) = \sigma(t)\sigma(t)'\pi_0(t) - \sigma(t)Z_\theta(t) = 0$, owing to the fact that $\pi_0(t)$ minimizes $|\sigma(t)'\pi - Z_\theta(t)|^2$. This implies that $\sigma(t)D(t) = 0$ and hence

$$\theta + D \in \Theta. \tag{4.12}$$

On the other hand, $Z_\theta(t)'\bar{D}(t) = [\pi_0(t)'\sigma(t)\bar{D}(t) - \bar{D}(t)'\bar{D}(t)] = -|\bar{D}(t)|^2$; hence $Z_\theta(t)'D(t) = -|\bar{D}(t)|$.

Define $\hat{X}(\cdot)$ to be the solution of the following (forward) SDE:

$$\begin{cases} d\hat{X}(t) = [r(t)\hat{X}(t) + (\theta(t) + D(t))'Z_\theta(t)]dt + Z_\theta(t)'dW(t), \\ \hat{X}(0) = x_0. \end{cases}$$

Ito's formula implies

$$E[\hat{X}(T)H_{\theta+D}(T)] = \hat{X}(0) = x_0 = E[X_\theta(T)H_{\theta+D}(T)], \tag{4.13}$$

where the last equality is due to (4.12) and the assumption. On the other hand,

$$d[\hat{X}(t) - X_\theta(t)] = r(t)[\hat{X}(t) - X_\theta(t)]dt + Z_\theta(t)'D(t)dt, \ \hat{X}(0) - X_\theta(0) = 0;$$

hence $\hat{X}(T) - X_\theta(T) = \int_0^T e^{\int_t^T r(s)ds} Z_\theta(t)'D(t)dt = -\int_0^T e^{\int_t^T r(s)ds}|\bar{D}(t)|dt$. Comparing this with (4.13) we conclude that $\bar{D}(t) = 0$, a.s., a.e.$t \in [0, T]$, which leads to (4.11). Since $\sigma(\cdot)'\pi_0(\cdot) = Z_\theta(\cdot) \in L^2_{\mathcal{F}}(0, T, \mathbf{R}^n)$, it follows that $\pi_0(\cdot) \in \Pi$. Now, the BSDE (4.10) that $(X_\theta(\cdot), Z_\theta(\cdot))$ satisfies can be rewritten as

$$\begin{cases} dX_\theta(t) = [r(t)X_\theta(t) + B(t)'\pi_0(t)]dt + \pi_0(t)'\sigma(t)dW(t), \\ X_\theta(T) = X, \end{cases}$$

which means that X is attained by the portfolio $\pi_0(\cdot)$. $\square$

Corollary 4.3.1 *A_C is a linear subspace of $L^2(\mathcal{F}_T, \mathbf{R})$.*

By Theorem 4.3.1, we can rewrite problem (4.8) as follows:

$$\begin{array}{ll} \text{Minimize} & E[X - (\lambda - \mu H_{\theta_0}(T))]^2, \\ \text{subject to} & \begin{cases} X \in L^2(\mathcal{F}_T, \mathbf{R}), \\ E[X(H_\theta(T) - H_{\theta_0}(T)] = 0, \ \ \forall \theta \in \Theta. \end{cases} \end{array} \tag{4.14}$$

Note that (4.14) is an optimization problem with *infinitely many* constraints. Moreover, since its objective function is strictly convex, there is at most one optimal solution. Denote

$L := \text{span}\{H_{\theta_1}(T) - H_{\theta_2}(T) : \theta_1 \in \Theta, \theta_2 \in \Theta\}$, where span$(A)$ means the minimal linear space that contains A, and consider $\bar{L}$, the closure of L in the $L^2(\mathcal{F}_T, \mathbf{R})$-norm. Since each $H_\theta(T) \in L^2(\mathcal{F}_T, \mathbf{R})$, it follows that $\bar{L} \subset L^2(\mathcal{F}_T, \mathbf{R})$. The following theorem provides a way to solving (4.14).

Theorem 4.3.2 *For any given (λ, μ), consider the following problem*

$$\begin{aligned} &\textit{Minimize} \quad E(\lambda - \mu H_{\theta_0}(T) - Y)^2, \\ &\textit{subject to} \quad Y \in \bar{L}. \end{aligned} \tag{4.15}$$

We have the following conclusion:

(i) *Problem (4.15) admits a unique optimal solution $Y^* \in \bar{L}$.*

(ii) *Y^* is the optimal solution to (4.15) if and only if Y^* is the only $Y \in \bar{L}$ such that $\lambda - \mu H_{\theta_0}(T) - Y \in A_C$.*

(iii) *The unique optimal solution to (4.14) can be expressed as $X^* = \lambda - \mu H_{\theta_0}(T) - Y^*$*

Proof: (i) First of all, $\bar{L}$ is a closed linear space. By the projection theorem in Hilbert spaces (refer to, e.g., [38, p.51, Theorem 2]), (4.15) has a unique optimal solution Y^*.

(ii) Y^* is optimal for (4.15) if and only if $E[(\lambda - \mu H_{\theta_0}(T) - Y^*)Y] = 0$, $\forall Y \in \bar{L}$. The latter implies that $X^* = \lambda - \mu H_{\theta_0}(T) - Y^*$ is feasible for (4.14). From Theorem 4.3.1, it follows that $\lambda - \mu H_{\theta_0}(T) - Y^* \in A_C$. By the uniqueness of the optimal solution for problem (4.15), we know Y^* is the only one in $\bar{L}$ such that $\lambda - \mu H_{\theta_0} - Y \in A_C$.

(iii) We have proved in (i) that $X^* = \lambda - \mu H_{\theta_0}(T) - Y^*$ is feasible for (4.14) if Y^* is optimal for (4.15). Now, for any feasible solution X of (4.14):

$$\begin{aligned} & E[X - (\lambda - \mu H_{\theta_0}(T))]^2 \\ = \; & E[X - (\lambda - \mu H_{\theta_0}(T)) + Y^*]^2 + 2E[Y^*(\lambda - \mu H_{\theta_0}(T))] - 2E[XY^*] - E[Y^*]^2 \\ = \; & E[X - (\lambda - \mu H_{\theta_0}(T) - Y^*)]^2 + 2E[Y^*(\lambda - \mu H_{\theta_0}(T))] - 2E[X^*Y^*] - E[Y^*]^2 \\ \geq \; & E[X^* - (\lambda - \mu H_{\theta_0}(T) - Y^*)]^2 + 2E[Y^*(\lambda - \mu H_{\theta_0}(T))] - 2E[X^*Y^*] - E[Y^*]^2 \\ = \; & E[X^* - (\lambda - \mu H_{\theta_0}(T))]^2, \end{aligned}$$

where we have used the fact that $E[XY^*] = E[X^*Y^*] = 0$ due to the constraint of problem (4.14). Hence X^* is the unique optimal solution to (4.14). $\square$

The above theorem suggests that one can obtain the optimal solution to (4.14), hence that to (4.8), via the (unique) optimal solution to the projection problem (4.15). In some special cases, especially in the case when the market parameters, $r(\cdot), \mu(\cdot)$ and $\sigma(\cdot)$, are all deterministic processes, (4.15) can be solved completely which in turn leads to the complete solution to the underlying mean–variance portfolio selection problem.

Lemma 4.3.1 *If $r(\cdot), \mu(\cdot)$ and $\sigma(\cdot)$ are deterministic, then $\lambda - \mu H_{\theta_0}(T) \in A_C$ for any (λ, μ).*

Proof: Fix $\theta(\cdot) \in \Theta$. By the definition of $\theta_0(\cdot)$, see (4.2), point-wisely $\theta_0(t)$ is the projection of 0 onto the closed affine space $\{\theta : \sigma(t)\theta = B(t)\} \subset \mathbf{R}^n$. Hence $\theta_0(t)'[\theta(t) - \theta_0(t)] = 0$ or $\theta_0(t)'\theta(t) = |\theta_0(t)|^2$. Now,

$$\begin{aligned}
& H_\theta(T)H_{\theta_0}(T) \\
= \;& \exp\{-\int_0^T [2r(t) + \frac{1}{2}(|\theta(t)|^2 + |\theta_0(t)|^2)]dt - \int_0^T [\theta(t) + \theta_0(t)]'dW(t)\} \\
= \;& \exp\{-\int_0^T [2r(t) - |\theta_0(t)|^2]dt\} \exp\{-\int_0^T \frac{1}{2}|\theta(t) + \theta_0(t)|^2 dt - \int_0^T [\theta(t) + \theta_0(t)]'dW(t)\}.
\end{aligned}$$

Thus, $E[H_\theta(T)H_{\theta_0}(T)] = \exp\{-\int_0^T [2r(t) - |\theta_0(t)|^2]dt\}$ which is independent of $\theta \in \Theta$. We have then $H_{\theta_0}(T) \in A_C$ thanks to Theorem 4.3.1. The conclusion follows as $\lambda \in A_C$. $\square$

Theorem 4.3.3 *If $r(\cdot), \mu(\cdot)$ and $\sigma(\cdot)$ are deterministic, then the efficient portfolio for the mean–variance problem (4.5) corresponding to z is the one that replicates the terminal claim $\lambda - \mu H_{\theta_0}(T)$, where*

$$\lambda = \frac{ze^{\int_0^T |\theta_0(t)|^2 dt} - x_0 e^{\int_0^T r(t)dt}}{e^{\int_0^t |\theta_0(t)|^2 dt} - 1}, \qquad \mu = \frac{ze^{\int_0^T r(t)dt} - x_0 e^{\int_0^T 2r(t)dt}}{e^{\int_0^T |\theta_0(t)|^2 dt} - 1}. \tag{4.16}$$

Moreover, the efficient frontier is

$$\mathrm{Var}(x(T)) = \frac{1}{e^{\int_0^T |\theta_0(t)|^2 dt} - 1}[z - x_0 e^{\int_0^T r(t)dt}]^2, \;\; z \geq x_0 e^{\int_0^T r(t)dt}. \tag{4.17}$$

Proof: By virtue of Lemma 4.3.1 and Theorem 4.3.2, $Y^* = 0$ is the unique optimal solution to (4.15), or $X^* = \lambda - \mu H_{\theta_0}(T)$ is the unique optimal solution to (4.14). To determine (λ, μ) so as to obtain the solution to (4.7), we apply Theorem 4.2.2 to derive the following system of equations

$$\begin{cases} \lambda - \mu E H_{\theta_0}(T) = z, \\ \lambda E H_{\theta_0}(T) - \mu E H_{\theta_0}(T)^2 = x_0. \end{cases}$$

Solving these equations we get

$$\lambda = \frac{z E H_{\theta_0}(T)^2 - x_0 E H_{\theta_0}(T)}{\mathrm{Var}(H_{\theta_0}(T))}, \qquad \mu = \frac{z E H_{\theta_0}(T) - x_0}{\mathrm{Var}(H_{\theta_0}(T))}.$$

Substituting $E H_{\theta_0}(T) = e^{-\int_0^T r(t)dt}$, $E H_{\theta_0}(T)^2 = e^{-\int_0^T [2r(t) - |\theta_0(t)|^2]dt}$ to the above, we get the expressions (4.16). Finally, the variance of the optimal terminal wealth is

$$\begin{aligned} \mathrm{Var}(x(T)) &= \mathrm{Var}(X^*) = \mu^2 \mathrm{Var}(H_{\theta_0}(T)) \\ &= \frac{[z E H_{\theta_0}(T) - x_0]^2}{\mathrm{Var}(H_{\theta_0}(T))} \\ &= \frac{1}{e^{\int_0^T |\theta_0(t)|^2 dt} - 1} [z - x_0 e^{\int_0^T r(t)dt}]^2. \end{aligned}$$

$\square$

4.4 Case 2: No-shorting case

Again, we need to first characterize the attainable set A_C in this case. Define $\hat{\Theta} := \{\theta \in L^\infty_{\mathcal{F}}(0, T, \mathbf{R}^n) : \sigma(t)\theta(t) \geq B(t),\ \text{a.s.},\ \text{a.e.} t \in [0, T]\}$, where the greater or equal relation between two vectors is in the component-wise sense.

Theorem 4.4.1 *For any $X \in L^2(\mathcal{F}_T, \mathbf{R})$, $X \in A_C$ if and only if there exists $\theta^* \in \hat{\Theta}$ such that $\sup_{\theta \in \hat{\Theta}} E[X H_\theta(T)] = E[X H_{\theta^*}(T)]$. Furthermore, $\sup_{\theta \in \hat{\Theta}} E[X H_\theta(T)] = E[X H_{\theta_0}(T)]$ if $X \in A_C$.*

Proof: If $X \in A_C$, then there is $(x(\cdot), \pi(\cdot)) \in C$ satisfying (2.5) with $x_0 = E[X H_{\theta_0}(T)]$. Take any $\theta \in \hat{\Theta}$ and consider $H_\theta(\cdot)$ that satisfies (4.4). Applying Ito's formula we get easily

$$d[x(t)H_\theta(t)] = [B(t) - \sigma(t)\theta(t)]'\pi(t)H_\theta(t)dt + [\pi(t)'\sigma(t) - x(t)\theta(t)']H_\theta(t)dW(t);$$

thus $E[X H_\theta(T)] = x_0 + E\int_0^T [B(t) - \sigma(t)\theta(t)]'\pi(t)H_\theta(t)dt \leq x_0 = E[X H_{\theta_0}(T)]$. Thus $\sup_{\theta \in \hat{\Theta}} E[X H_\theta(T)] = E[X H_{\theta_0}(T)]$.

Conversely, suppose there is $\theta^* \in \hat{\Theta}$ such that $x_0 := E[XH_{\theta^*}(T)] \geq E[XH_\theta(T)]$ $\forall\theta \in \hat{\Theta}$. Let $(X^*(\cdot), Z^*(\cdot))$ be the unique solution to the following BSDE

$$\begin{cases} dX^*(t) = [r(t)X^*(t) + \theta^*(t)'Z^*(t)]dt + Z^*(t)'dW(t), \\ X^*(T) = X. \end{cases} \tag{4.18}$$

We are to show that there exist an no-shorting admissible portfolio $\pi_0(\cdot)$ such that

$$Z^*(t) = \sigma(t)'\pi_0(t), \text{ a.s., a.e.} t \in [0,T]. \tag{4.19}$$

Indeed, define

$$\pi_0(t) := \text{argmin}_{\pi\in\text{argmin}_{\pi\in\mathbf{R}^m_+}|\sigma(t)'\pi - Z^*(t)|^2}|\pi|^2.$$

Again, thanks to Lemma A.2, $\pi_0(\cdot)$ is a progressively measurable stochastic processes with respect to $\mathcal{F}_t$. Note $Z^*(t) \notin \{\sigma(t)'\pi : \pi \in \mathbf{R}^m_+\}$ whenever $\sigma(t)'\pi_0(t) - Z^*(t) \neq 0$. Thus by Lemma B.3 and A.2, there is $\bar{D}(\cdot)$ which is progressively measurable satisfying $\bar{D}(t) \neq 0, Z^*(t)'\bar{D}(t) < 0, \sigma(t)\bar{D}(t) \geq 0$, a.s., a.e.$t$, on the set where $\sigma(t)'\pi_0(t) - Z^*(t) \neq 0$. Set

$$D(t) := \begin{cases} 0, & \text{if } \sigma(t)'\pi_0(t) - Z^*(t) = 0, \\ \bar{D}(t)/|\bar{D}(t)|, & \text{if } \sigma(t)'\pi_0(t) - Z^*(t) \neq 0. \end{cases}$$

Then $D(\cdot) \in L^\infty_{\mathcal{F}}(0,T,\mathbf{R}^n)$, $\sigma(t)D(t) \geq 0$, and

$$Z^*(t)'D(t) < 0 \quad \text{whenever} \quad \sigma(t)'\pi_0(t) - Z^*(t) \neq 0. \tag{4.20}$$

Since $\sigma(t)[\theta^*(t) + D(t)] \geq \sigma(t)\theta^*(t) \geq B(t)$, we conclude $\theta^* + D \in \hat{\Theta}$.

Define $\bar{X}(\cdot)$ to be the solution of the following SDE:

$$\begin{cases} d\bar{X}(t) = [r(t)\bar{X}(t) + (\theta^*(t) + D(t))'Z^*(t)]dt + Z^*(t)'dW(t), \\ \bar{X}(0) = x_0. \end{cases}$$

Then

$$E[\bar{X}(T)H_{\theta^*+D}(T)] = \bar{X}(0) = x_0 = E[X^*(T)H_{\theta^*}(T)] \geq E[X^*(T)H_{\theta^*+D}(T)]. \tag{4.21}$$

On the other hand,

$$d[\bar{X}(t) - X^*(t)] = r(t)[\bar{X}(t) - X^*(t)]dt + Z^*(t)'D(t)dt, \ \bar{X}(0) - X^*(0) = 0;$$

hence $\bar{X}(T) - X^*(T) = \int_0^T e^{\int_t^T r(s)ds} Z^*(t)'D(t)dt$. It then follows from (4.21) and (4.20) that $\sigma(t)'\pi_0(t) - Z^*(t) = 0$, a.s., a.e.$t \in [0,T]$. This proves (4.19).

Next, let $\hat{X}(\cdot)$ be the solution to the following SDE:

$$\begin{cases} d\hat{X}(t) = [r(t)\hat{X}(t) + \theta_0(t)'Z^*(t)]dt + Z^*(t)'dW(t), \\ \hat{X}(t) = x_0. \end{cases} \tag{4.22}$$

Then $E[\hat{X}(T)H_{\theta_0}(T)] = x_0 \geq E[X^*(T)H_{\theta_0}(T)]$. On the other hand,

$$\begin{aligned} d[\hat{X}(t) - X^*(t)] = & \quad r(t)[\hat{X}(t) - X^*(t)]dt + [\theta_0(t) - \theta^*(t)]'Z^*(t)dt \\ = & \quad r(t)[\hat{X}(t) - X^*(t)]dt + [B(t) - \sigma(t)\theta^*(t)]'\pi_0(t)dt, \end{aligned}$$

where we have used the fact that $Z^*(t) = \sigma(t)'\pi_0(t)$. Hence,

$$\hat{X}(T) - X^*(T) = \int_0^T e^{\int_t^T r(s)ds}[B(t) - \sigma(t)\theta^*(t)]'\pi_0(t)dt \leq 0$$

By $E[(\hat{X}(T) - X^*(T))H_{\theta_0}(T)] \geq 0$ we have

$$B(t)'\pi_0(t) = \theta^*(t)'\sigma(t)'\pi_0(t) \equiv \theta^*(t)'Z^*(t), \tag{4.23}$$

and

$$E[XH_{\theta^*}(T)] \equiv x_0 \equiv E[\hat{X}(T)H_{\theta_0}(T)] = E[XH_{\theta_0}(T)]. \tag{4.24}$$

It follows from (4.18) and (4.23) that $(X^*(\cdot), \pi_0(\cdot))$ satisfies

$$\begin{cases} dX^*(t) = [r(t)X^*(t) + B(t)'\pi_0(t)]dt + \pi_0(t)'\sigma(t)dW(t), \\ X^*(T) = X, \end{cases}$$

meaning that $X \in A_C$. Finally, the second assertion of the theorem follows from (4.24).
□

By virtue of Theorem 4.4.1, the problem (4.8) for Case 2 can be written as

$$\begin{array}{ll} \text{Minimize} & E[X - (\lambda - \mu H_{\theta_0}(T))]^2, \\ \text{subject to} & \begin{cases} X \in L^2(\mathcal{F}_T, \mathbf{R}), \\ \max_{\theta \in \hat{\Theta}} E[XH_\theta(T)] = E[XH_{\theta_0}(T)]. \end{cases} \end{array} \tag{4.25}$$

Denote $M := \overline{\text{cone}\{H_\theta(T) - H_{\theta_0}(T) : \theta \in \hat{\Theta}\}}$. which can be easily verified to be a closed convex cone.

The following theorem is the no-shorting counterpart of Theorem 4.3.2.

Theorem 4.4.2 *For any given (λ, μ), consider the following problem*

$$\begin{aligned} &\textit{Minimize} \quad E(\lambda - \mu H_{\theta_0}(T) - Y)^2, \\ &\textit{subject to} \quad Y \in M. \end{aligned} \tag{4.26}$$

We have the following conclusion:

(i) *Y^* is the optimal solution to (4.26) if and only if $Y^* \in M$ and*

$$\lambda - \mu H_{\theta_0}(T) - Y^* \in A_C, \; E[(\lambda - \mu H_{\theta_0}(T) - Y^*)Y^*] = 0. \tag{4.27}$$

(ii) *The unique optimal solution to (4.25) can be expressed as $X^* = \lambda - \mu H_{\theta_0}(T) - Y^*$, where Y^* is the unique optimal solution to (4.26).*

Proof: (i) First of all, (4.26) is an optimization problem with a coercive, strictly convex cost function and a nonempty, closed convex constraint set, which therefore must admit a unique optimal solution. Moreover, Y^* is optimal to (4.26) if and only if $Y^* \in M$ and for any $Y \in M$, $0 \in \text{argmin}_{0 \le \alpha \le 1} E[f(\alpha)]$ where $f(\alpha) := [\lambda - \mu H_{\theta_0}(T) - Y^* + \alpha(Y^* - Y)]^2$.

Denote $X^* = \lambda - \mu H_{\theta_0}(T) - Y^*$, then

$$\begin{aligned} Ef(\alpha) - Ef(0) &= E[(2X^* + \alpha(Y^* - Y))(\alpha(Y^* - Y))] \\ &= 2\alpha E[X^*(Y^* - Y)] + \alpha^2 E(Y^* - Y)^2. \end{aligned}$$

so $0 \in \text{argmin}_{0 \le \alpha \le 1} E[f(\alpha)]$ if and only if $2E[X^*(Y^* - Y)] + \alpha E(Y^* - Y)^2 \ge 0$ for any $\alpha \in [0, 1]$, which is equivalent with $E[X^*(Y^* - Y)] \ge 0$. Therefore, Y^* is optimal if and only if $Y^* \in M$ and

$$E[(\lambda - \mu H_{\theta_0}(T) - Y^*)(Y^* - Y)] \ge 0 \;\; \forall Y \in M. \tag{4.28}$$

To prove that (4.27) and (4.28) are equivalent, first note that (4.27) easily yields (4.28) thanks to Theorem 4.4.1. Now, suppose (4.28) holds. Taking $Y = 0 \in M$ we get from (4.28) that $E[(\lambda - \mu H_{\theta_0}(T) - Y^*)Y^*] \ge 0$, and taking $Y = 2Y^* \in M$ (recall that M is a cone) we get $E[(\lambda - \mu H_{\theta_0}(T) - Y^*)Y^*] \le 0$. Consequently $E[(\lambda - \mu H_{\theta_0}(T) - Y^*)Y^*] = 0$ and, together with (4.28), results in (4.27).

(ii) We proved in (i) that $X^* = \lambda - \mu H_{\theta_0}(T) - Y^*$ is feasible for (4.25) if Y^* is optimal for (4.26). On the other hand, for any feasible solution X of (4.25):

$$E[X - (\lambda - \mu H_{\theta_0}(T))]^2$$

$$
\begin{aligned}
&= E[X-(\lambda-\mu H_{\theta_0}(T))+Y^*]^2+2E[Y^*(\lambda-\mu H_{\theta_0}(T))]-2E[XY^*]-E[Y^*]^2\\
&\geq E[X-(\lambda-\mu H_{\theta_0}(T)-Y^*)]^2+2E[Y^*(\lambda-\mu H_{\theta_0}(T))]-2E[X^*Y^*]-E[Y^*]^2\\
&\geq E[X^*-(\lambda-\mu H_{\theta_0}(T)-Y^*)]^2+2E[Y^*(\lambda-\mu H_{\theta_0}(T))]-2E[X^*Y^*]-E[Y^*]^2\\
&= E[X^*-(\lambda-\mu H_{\theta_0}(T))]^2,
\end{aligned}
$$

where we have used the fact that $E[XY^*] \leq 0$ and $E[X^*Y^*] = 0$. This means that X^* is an optimal solution for (4.25). □

As with Case 1, we now discuss the case when all the market coefficients are deterministic and show how to apply Theorem 4.4.2 to solve the mean-variance problem.

Lemma 4.4.1 *If $r(\cdot), B(\cdot)$ and $\sigma(\cdot)$ are deterministic, then for any given (λ, μ) with $\mu \geq 0$, (4.26) has the optimal solution $Y^* := \mu(H_{\theta_1^*}(T) - H_{\theta_0}(T))$, where*

$$\theta_1^*(t) := \text{argmin}_{\theta\in\{\theta\in\mathbf{R}^n:\sigma(t)\theta\geq B(t)\}}|\theta|^2. \tag{4.29}$$

Proof: First we prove that

$$\theta_0(t)'\theta_1^*(t) = |\theta_1^*(t)|^2, \quad \forall t \in [0,T]. \tag{4.30}$$

Indeed, $\theta_1^*(t)$ is well-defined. And by the Largrange method, for any $t \in [0,T]$, there exists a $k(t) \in \mathbf{R}_+^m$ such that $\theta_1^*(t) = \sigma(t)'k(t)$ and $k(t)'(\sigma(t)\theta_1^*(t) - B(t)) = 0$. so

$$\theta_0(t)'\theta_1^*(t) = \theta_0(t)'\sigma(t)'k(t) = k(t)'B(t) = k(t)'\sigma(t)\theta_1^*(t) = |\theta_1^*(t)|^2.$$

Next, according to Theorem 4.4.2, to prove that $Y^* := \mu(H_{\theta_1}(T) - H_{\theta_0}(T))$ is optimal to (4.26) it suffices to show that

$$
\begin{aligned}
&E[(\lambda-\mu H_{\theta_1^*}(T))(H_{\theta_1^*}(T)-H_{\theta_0}(T))] = 0,\\
&E[(\lambda-\mu H_{\theta_1^*}(T))(H_{\theta}(T)-H_{\theta_0}(T))] \leq 0, \quad \forall\theta\in\hat{\Theta}.
\end{aligned}
$$

Since in the present case of deterministic coefficients the value of $E[H_\theta(T)]$ is independent of $\theta \in \hat{\Theta}$, the above is equivalent to

$$E[H_{\theta_1^*}(T)(H_{\theta_1^*}(T)-H_{\theta_0}(T))] = 0, \quad \text{and } E[H_{\theta_1^*}(T)(H_\theta(T)-H_{\theta_0}(T))] \geq 0 \ \ \forall\theta\in\hat{\Theta}. \tag{4.31}$$

Now,

$$H_{\theta_1^*}(T)H_{\theta_0}(T)$$

$$
\begin{aligned}
&= \exp\{-\int_0^T [2r(t) + \frac{1}{2}(|\theta_1^*(t)|^2 + |\theta_0(t)|^2)]dt - \int_0^T [\theta_1^*(t) + \theta_0(t)]' dW(t)\} \\
&= \exp\{-\int_0^T [2r(t) - \theta_1^*(t)'\theta_0(t)]dt\} \exp\{-\int_0^T \frac{1}{2}|\theta(t) + \theta_0(t)|^2 dt - \int_0^T [\theta(t) + \theta_0(t)]' dW(t)\} \\
&= \exp\{-\int_0^T [2r(t) - |\theta_1^*(t)|^2]dt\} \exp\{-\int_0^T \frac{1}{2}|\theta(t) + \theta_0(t)|^2 dt - \int_0^T [\theta(t) + \theta_0(t)]' dW(t)\}.
\end{aligned}
$$

Thus, $E[H_{\theta_1^*}(T)H_{\theta_0}(T)] = \exp\{-\int_0^T [2r(t) - |\theta_1^*(t)|^2]dt\} = E[H_{\theta_1^*}(T)^2]$, which proves the first equality of (4.31). Next, the definition of θ_1^* ensures that $\theta(t)'\theta_1^*(t) \geq |\theta_1^*(t)|^2 \quad \forall t \in [0,T], \forall \theta \in \hat{\Theta}$. Now, a similar calculation as above shows that, for any $\theta \in \hat{\Theta}$, $E[H_{\theta_1^*}(T)H_\theta(T)] = \exp\{-\int_0^T [2r(t) - \theta_1^*(t)'\theta(t)]dt\} \geq \exp\{-\int_0^T [2r(t) - |\theta_1^*(t)|^2]dt\} = E[H_{\theta_1^*}(T)^2]$. This, together with the proved first equality of (4.31), leads to the second inequality of (4.31). □

Theorem 4.4.3 *If $r(\cdot)$, $B(\cdot)$ and $\sigma(\cdot)$ are deterministic, and $\int_0^T \sum_{i=1}^n B(t)_i^+ dt > 0$, then the efficient portfolio for the mean–variance problem (4.5) corresponding to $z \geq x_0 e^{\int_0^T r(s)ds}$ is the one that replicates the terminal claim $\lambda - \mu H_{\theta_1^*}(T)$, where*

$$
\lambda = \frac{z e^{\int_0^T |\theta_1^*(t)|^2 dt} - x_0 e^{\int_0^T r(t)dt}}{e^{\int_0^t |\theta_1^*(t)|^2 dt} - 1}, \qquad \mu = \frac{z e^{\int_0^T r(t)dt} - x_0 e^{\int_0^T 2r(t)dt}}{e^{\int_0^T |\theta_1^*(t)|^2 dt} - 1}. \tag{4.32}
$$

Moreover, the efficient frontier is

$$
\mathrm{Var}(x(T)) = \frac{1}{e^{\int_0^T |\theta_1^*(t)|^2 dt} - 1}[z - x_0 e^{\int_0^T r(t)dt}]^2, \quad z \geq x_0 e^{\int_0^T r(t)dt}. \tag{4.33}
$$

Proof: By virtue of Lemma 4.4.1 and Theorem 4.4.2, $X^* := \lambda - \mu H_{\theta_1}(T)$ is the unique optimal solution to (4.25), provided that $\mu \geq 0$. The system of equations (4.9) reduces to

$$
\begin{cases}
\lambda - \mu E H_{\theta_1}(T) = z \\
\lambda E H_{\theta_1}(T) - \mu E H_{\theta_1}(T)^2 = x_0,
\end{cases}
$$

where we have used the fact that $E[H_{\theta_1}(T)H_{\theta_0}(T)] = E[H_{\theta_1}(T)^2]$ which was proved in the proof of Lemma 4.4.1. Clearly (4.32) gives the (only) solution pair to the above system with $\mu \geq 0$ under the assumption that $z \geq x_0 e^{\int_0^T r(t)dt}$. Finally, the efficient frontier (4.33) can be derived in exactly the same way as in the unconstrained case; see the proof of Theorem 4.3.3. □

Remark 4.4.1 The condition $\int_0^T \sum_{i=1}^n B(t)_i^+ ds > 0$ is mild. It does not hold only when $B(t) \leq 0$ for $a.e.t \in [0,T]$, which means the appreciation rates of all the risky assets are no higher than the interest rate, which is unreasonable.

4.5 Case 3: No bankruptcy case

In Section 3.2 we have shown that in a complete market, a wealth process $x(\cdot) \geq 0$ if and only if its terminal wealth $x(T) \geq 0$. This conclusion still holds, due to Assumption 4.2.1, in the incomplete market specified in Section 4.2. Hence in Case 3 the constraint set can be written as $C = \{(x(\cdot), \pi(\cdot)) \in L^2_{\mathcal{F}}(0, T, \mathbf{R}) \times \Pi : x(T) \geq 0\}$.

The following result can be proved in exactly the same way as that of Theorem 4.3.1.

Theorem 4.5.1 *For any* $X \in L^2(\mathcal{F}_T, \mathbf{R})$, $X \in A_C$ *if and only if* $X \geq 0$ *and* $E[XH_\theta(T)]$ *is independent of* $\theta(\cdot) \in \Theta_1 \stackrel{\triangle}{=} \{\theta(\cdot) : \sigma(t)\theta(t) = B(t), |\theta(t) - \theta_0(t)| \leq 1,\ a.s.,\ a.e. t \in [0, T]\}$.

In view of Theorem 4.5.1, we can rewrite problem (4.8) for Case 3 as follows:

$$
\begin{array}{ll}
\text{Minimize} & E[X - (\lambda - \mu H_{\theta_0}(T))]^2, \\
\text{subject to} & \left\{ \begin{array}{l} X \in L^2(\mathcal{F}_T, \mathbf{R}), \\ \sup_{\theta_1, \theta_2 \in \Theta_1} E[X(H_{\theta_1}(T) - H_{\theta_2}(T))] \leq 0, \\ X \geq 0. \end{array} \right.
\end{array}
\tag{4.34}
$$

Denote $M_1 := \overline{\{H_{\theta_1}(T) - H_{\theta_2}(T) : \theta_1, \theta_2 \in \Theta_1\}}$. It is clear that $M_1 \subset L^2(\mathcal{F}_T, \mathbf{R})$, and M_1 is a bounded convex set in view of Lemma 4.2.1.

Lemma 4.5.1 *For any* $0 \leq k \leq 1$, *if* $Y \in M_1$, *then* $kY \in M_1$.

Proof: For any $Y = H_{\theta_1}(T) - H_{\theta_2}(T)$ with $\theta_1 \in \Theta_1, \theta_2 \in \Theta_1$,

$$
\begin{aligned}
kY &= kH_{\theta_1}(T) - kH_{\theta_2}(T) \\
&= kH_{\theta_1}(T) + (1-k)H_{\theta_2}(T) - H_{\theta_2}(T) \\
&= H_{\theta_3}(T) - H_{\theta_2}(T) \\
&\in M.
\end{aligned}
$$

Next, for any $Y \in M_1$, there exists a sequence $Y_n \in \{H_{\theta_1}(T) - H_{\theta_2}(T) : \theta_1 \in \Theta_1, \theta_2 \in \Theta_1\}$ such that $Y_n \to Y$. Therefore $kY_n \in M$ and $kY_n \to kY$, which implies $kY \in M_1$. $\square$

It is easy to see that problem (4.34) is equivalent to

$$\begin{aligned} &\text{Minimize} \quad E[X - (\lambda - \mu H_{\theta_0}(T))]^2, \\ &\text{subject to} \quad \begin{cases} X \in L^2(\mathcal{F}_T, \mathbf{R}), \\ \sup_{Y \in M_1} E[XY] \le 0, \\ X \ge 0, \end{cases} \end{aligned} \tag{4.35}$$

which admits a unique optimal solution. For written convenience, denote $\hat{\lambda} = \lambda - \mu H_{\theta_0}(T)$.

Theorem 4.5.2 *Let $F(\cdot)$ be a strictly convex, coercive, and lower bounded function from $L^2(\mathcal{F}_T, \mathbf{R})$ to $\mathbf{R}$, $G(\cdot)$ a convex function from $L^2(\mathcal{F}_T, \mathbf{R})$ to $\mathbf{R}$, and $\{X : G(X) \le 0\}$ nonempty. If X^* is the optimal solution for*

$$\begin{aligned} &\text{Minimize} \quad F(X), \\ &\text{subject to} \quad G(X) \le 0, \end{aligned} \tag{4.36}$$

then there exists $k \ge 0$ such that X^* is also the unique solution for

$$\min_{X \in L^2(\mathcal{F}_T, \mathbf{R})} \{F(X) + kG(X)\}. \tag{4.37}$$

Proof: By the properties of $F(\cdot)$, we know problem (4.36) admits a unique optimal solution X^*. Define $A \triangleq \{a : \{X \in L^2(\mathcal{F}_T, \mathbf{R}) : G(X) \le a\} \neq \varnothing\}$. Then A is a convex set, and for any $a \in A$, $f(a) = \min_{X \in L^2(\mathcal{F}_T, \mathbf{R}), G(X) \le a} F(X)$ is well defined and bounded by some real number. Obviously, $f(a)$ is decreasing on $\mathbf{R}^+$. Now we claim $f(a)$ is convex. Indeed, for any $l \in [0,1]$, $a \in A, b \in A$,

$$\begin{aligned} f(la + (1-l)b) &= \min_{X \in L^2(\mathcal{F}_T, \mathbf{R}), G(X) \le la + (1-l)b} F(X) \\ &= \min_{(lX_1 + (1-l)X_2 \in L^2(\mathcal{F}_T, \mathbf{R}), G(lX_1 + (1-l)X_2) \le la + (1-l)b} F(lX_1 + (1-l)X_2) \\ &\le \min_{(X_1 \in L^2(\mathcal{F}_T, \mathbf{R}), X_2 \in L^2(\mathcal{F}_T, \mathbf{R}), G(X_1) \le a, G(X_2) \le b} F(lX_1 + (1-l)X_2) \\ &\le \min_{(X_1 \in L^2(\mathcal{F}_T, \mathbf{R}), X_2 \in L^2(\mathcal{F}_T, \mathbf{R}), G(X_1) \le a, G(X_2) \le b} lF(X_1) + (1-l)F(X_2) \\ &= l \min_{X_1 \in L^2(\mathcal{F}_T, \mathbf{R}), G(X_1) \le a} F(X_1) + (1-l) \min_{X_2 \in L^2(\mathcal{F}_T, \mathbf{R}), G(X_2) \le b} F(X_2) \\ &= lf(a) + (1-l)f(b). \end{aligned}$$

Now take $-k$ as the right hand side derivative of f at 0, then $k \ge 0$, and by the convexity of $f(\cdot)$, we have $f(a) - f(0) \ge -ka$ for any $a \in A$, which implies $f(a) + ka \ge f(0)$.

Therefore for any $X \in L^2(\mathcal{F}_T, \mathbf{R})$, denoting $a = G(X) \in A$, we have

$$F(X) + kG(X) \geq f(a) + ka \geq f(0) \geq F(X^*) + kG(X^*).$$

This shows X^* is optimal for problem (4.37); so k meets the requirement of the theorem. □

Lemma 4.5.2 *Let X^* be the optimal solution to (4.35). Then there exist $k_0 \geq 0$ such that for any $k \geq k_0$, X^* is also the unique optimal solution for*

$$\min_{X \in L^2(\mathcal{F}_T, \mathbf{R}^+)} \{E(X - \hat{\lambda})^2 + 2k \sup_{Y \in M_1} E[XY]\}, \tag{4.38}$$

and the optimal value of (4.38) is equal to $E(X^* - \hat{\lambda})^2$.

Proof: Firstly, by Theorem 4.5.2, we know there exists $k_0 \geq 0$ such that X^* is the unique optimal solution for

$$\min_{X \in L^2(\mathcal{F}_T, \mathbf{R}^+)} \{E(X - \hat{\lambda})^2 + 2k_0 \sup_{Y \in M_1} E[XY]\}. \tag{4.39}$$

For any $k \geq k_0$, any $X \in L^2(\mathcal{F}_T, \mathbf{R}^+)$,

$$\begin{aligned} & E(X - \hat{\lambda})^2 + 2k \sup_{Y \in M_1} E[XY] \\ \geq\ & E(X - \hat{\lambda})^2 + 2k_0 \sup_{Y \in M_1} E[XY] \\ \geq\ & E(X^* - \hat{\lambda})^2 + 2k_0 \sup_{Y \in M_1} E[X^*Y] \\ =\ & E(X^* - \hat{\lambda})^2 + 2k \sup_{Y \in M_1} E[X^*Y]. \end{aligned}$$

So X^* is also the unique optimal for problem (4.38) for any $k \geq k_0$. □

Lemma 4.5.3 *For any $k \geq k_0$, the unique optimal solution for problem (4.38) is $X_k^* = (\hat{\lambda} - kY_k^*)^+$, where Y_k^* is the optimal solution for*

$$\min_{Y \in M_1} E[(\hat{\lambda} - kY)^+]^2. \tag{4.40}$$

Proof: First of all, problem (4.40) admits optimal solutions. Rewrite problem (4.38) as

$$\min_{X \in L^2(\mathcal{F}_T, \mathbf{R}^+)} \sup_{Y \in M_1} \{E(X - \hat{\lambda})^2 + 2kE[XY]\}. \tag{4.41}$$

It is obvious that $f_k(X,Y) := E(X-\hat{\lambda})^2 + 2kE[XY]$ is convex in X and concave in Y. Furthermore, $L^2(\mathcal{F}_T, \mathbf{R}^+)$ is closed and convex, M_1 is bounded, closed, and convex. By Theorem 2.8.1 in [3](Page 55), we know for any optimal solution of problem (4.40), Y_k^*,

$$\begin{aligned}
& \min_{X\in L^2(\mathcal{F}_T,\mathbf{R}^+)} \sup_{Y\in M_1} \{E(X-\hat{\lambda})^2 + 2kE[XY]\} \\
= \quad & \sup_{Y\in M_1} \min_{X\in L^2(\mathcal{F}_T,\mathbf{R}^+)} \{E(X-(\hat{\lambda}-kY))^2 - E(\hat{\lambda}-kY)^2\} + E\hat{\lambda}^2 \\
= \quad & \sup_{Y\in M_1} \{E[(\hat{\lambda}-kY)^-]^2 - E(\hat{\lambda}-kY)^2\} + E\hat{\lambda}^2 \\
= \quad & -\inf_{Y\in M_1} E[(\hat{\lambda}-kY)^+]^2 + E\hat{\lambda}^2 \\
= \quad & -E[(\hat{\lambda}-kY_k^*)^+]^2 + E\hat{\lambda}^2.
\end{aligned}$$

By Lemma 4.5.2, the optimal value is independent of $k \geq k_0$. So $E[(\hat{\lambda}-kY_k^*)^+]^2 = E[(\hat{\lambda}-k_0Y_{k_0}^*)^+]^2$. By Lemma 4.5.1, for any $Y \in M_1, k' \geq 0$,

$$\begin{aligned}
E[(\hat{\lambda}-k'Y)^+]^2 \quad & \geq \quad E[(\hat{\lambda}-k'Y_{k'}^*)^+]^2 \\
& \geq \quad \min_{Y\in M_1} E[(\hat{\lambda}-(k'+k)Y)^+]^2 \\
& = \quad E[(\hat{\lambda}-(k+k')Y_{k+k'}^*)^+]^2 \\
& = \quad E[(\hat{\lambda}-kY_k^*)^+]^2.
\end{aligned}$$

Therefore $E[(\hat{\lambda}-kY_k^*)^+]^2 \leq E[(\hat{\lambda}-k'Y_k^*)^+]^2$ for any $k' \geq 0$. This implies $E[(\hat{\lambda}-kY_k^*)Y_k^*] = 0$. Therefore

$$\begin{aligned}
\sup_{Y\in M_1} \{E(X_k^*-\hat{\lambda})^2 + 2kE[X_k^*Y]\} \quad & = \quad E(X_k^*-\hat{\lambda})^2 + 2kEX_k^*Y_k^* \\
& = \quad -E[(\hat{\lambda}-kY_k^*)^+]^2 + E\hat{\lambda}^2 \\
& = \quad \min_{X\in L^2(\mathcal{F}_T,\mathbf{R}^+)} \sup_{Y\in M_1} \{E(X-\hat{\lambda})^2 + 2kE[XY]\}.
\end{aligned}$$

So X_k^* is optimal for problem (4.38). By the uniqueness of the optimal solution to (4.38), X_k^* is the only optimal solution. □

Corollary 4.5.1 *There exists $k_0 \geq 0$ such that the optimal solution X^* for problem (4.35) must be $X_k^* = (\hat{\lambda}-k_0Y_{k_0}^*)^+$, where $Y_{k_0}^*$ is the optimal solution for problem (4.39).*

Theorem 4.5.3 *Problem*

$$\min_{Y\in \text{cone}\{M_1\}} E[(\hat{\lambda}-Y)^+]^2 \tag{4.42}$$

admits optimal solutions in $L^2(\mathcal{F}_T, \mathbf{R})$*, and for different optimal solutions (if exist),* $(\hat{\lambda} - Y^*)^+$ *is the same, which is the unique optimal solution for problem (4.35) and problem (4.34). Furthermore,* Y^* *is optimal for problem (4.42) if and only if* $X^* = (\hat{\lambda} - Y^*)^+ \in A_C$.

In the remainder of this section, we study the case when all the market paramters are deterministic.

Lemma 4.5.4 *If* $r(\cdot), \mu(\cdot)$ *and* $\sigma(\cdot)$ *are deterministic, then* $[\lambda - \mu H_{\theta_0}(T)]^+ \in A_C$ *for any* (λ, μ).

Proof. For any $\theta \in \Theta$, by Girsanov's theorem $\tilde{W}(t) := W(t) + \int_0^t \theta(s)ds$ is a standard Brownian motion under a new probability measure $\tilde{P}$ defined as $d\tilde{P} = e^{\int_0^T r(t)dt} H_\theta(T)dP$. Now,

$$
\begin{aligned}
& E[H_\theta(T)[\lambda - \mu H_{\theta_0}(T)]^+] \\
= \; & E[H_\theta(T)[\lambda - \mu e^{-\int_0^T r(t)dt} e^{-\int_0^T [|\theta_0(t)|^2/2 - \theta_0(t)'\theta(t)]dt - \int_0^T \theta_0(t)' d\tilde{W}(t)}]^+] \\
= \; & e^{-\int_0^T r(t)dt} \tilde{E}[\lambda - \mu e^{-\int_0^T r(t)dt} e^{-\int_0^T [|\theta_0(t)|^2/2 - |\theta_0(t)|^2]dt - \int_0^T \theta_0(t)' d\tilde{W}(t)}]^+ \\
= \; & e^{-\int_0^T r(t)dt} E[\lambda - \mu e^{-\int_0^T [r(t) - |\theta_0(t)|^2/2]dt - \int_0^T \theta_0(t)' dW(t)}]^+,
\end{aligned}
$$

where we have used the fact that $\theta_0(t)'\theta(t) = |\theta_0(t)|^2$. This completes the proof in view of Theorem 4.5.1. $\Box$.

Define

$$
\begin{aligned}
& a := \inf\{\eta \in \mathbf{R} : P(H_{\theta_0}(T) < \eta) > 0\}, \\
& b = \sup\{\eta \in \mathbf{R} : P(H_{\theta_0}(T) > \eta) > 0\}.
\end{aligned}
\tag{4.43}
$$

Theorem 4.5.4 *If* $r(\cdot), \mu(\cdot)$ *and* $\sigma(\cdot)$ *are deterministic, and* $a < \frac{x_0}{z} < b$*, then the efficient portfolio for the mean–variance problem (4.5) corresponding to* z *is the one that replicates the terminal claim* $[\lambda - \mu H_{\theta_0}(T)]^+$*, where* (λ, μ) *is the unique solution to*

$$
\begin{cases}
E(\lambda - \mu H_{\theta_0}(T))^+ = z, \\
E[H_{\theta_0}(T)(\lambda - \mu H_{\theta_0}(T))^+] = x_0.
\end{cases}
\tag{4.44}
$$

Proof: By virtue of Lemma 4.5.4 and Theorem 4.5.3, $Y^* = 0$ is an optimal solution to (4.42), or $X^* = \lambda - \mu H_{\theta_0}(T)$ is the unique optimal solution to (4.34). To determine (λ, μ) so as to obtain the solution to (4.7), we apply Theorem 4.2.2 to derive the system of equations (4.44). The existence and uniqueness of the solution to (4.44) were established in Section 3.3. □

Remark 4.5.1 If we assume further that $\int_0^T |\theta_0(t)|^2 dt > 0$, then explicit forms of the efficient portfolios and efficient frontier can be obtained in exactly the same way as in Section 3.7 (by simply replacing $\rho(T)$ therein by $H_{\theta_0}(T)$).

4.6 Case 4: Neither shorting nor bankruptcy case

Similarly with Case 3, in Case 4 the constraint set can be written as $C = \{(x(\cdot), \pi(\cdot)) \in L^2_{\mathcal{F}}(0, T, \mathbf{R}) \times \Pi : x(T) \geq 0\}$.

Define $\theta_0(t) = \text{argmin}_{\sigma(t)x=B(t)}|x|^2$. Let K be a real number satisfying $|\theta_0(t)| \leq K$ *a.s.*, *a.e.*$t \in [0, t]$, $\hat{\Theta}_1 \triangleq \{\theta \in L^{+\infty}_{\mathcal{F}}(0, T, \mathbf{R}) : \sigma(\cdot)\theta(\cdot) \geq B, \exists N < K+1$ such that $|\theta(s, \omega)| \leq N$, *a.s.*, *a.e.*$t \in [0, T]\}$. It is obvious that $\hat{\Theta}_1$ is a nonempty convex set, and the set $\{H_\theta(T) : \theta \in \hat{\Theta}_1\}$ is bounded in L^2-norm.

Theorem 4.6.1 *For any $X \in L^2(\mathcal{F}_T, \mathbf{R})$, $X \in A_C$ if and only if $X \geq 0$ and there exists $\theta^* \in \hat{\Theta}_1$ such that $\sup_{\theta \in \hat{\Theta}_1} E[XH_\theta(T)] = E[XH_{\theta^*}(T)]$. Furthermore, $\sup_{\theta \in \hat{\Theta}_1} E[XH_\theta(T)] = E[XH_{\theta_0}(T)]$ if $X \in A_C$.*

Proof: The "if" part is the same as Theorem 4.4.1.

Conversely, suppose there is $\theta^* \in \hat{\Theta}_1$ such that $x_0 := E[XH_{\theta^*}(T)] \geq E[XH_\theta(T)]$ $\forall \theta \in \hat{\Theta}_1$. Let $(X^*(\cdot), Z^*(\cdot))$ be the unique solution to the following BSDE

$$\begin{cases} dX^*(t) = [r(t)X^*(t) + \theta^*(t)'Z^*(t)]dt + Z^*(t)'dW(t), \\ X^*(T) = X. \end{cases} \tag{4.45}$$

We are to show that there exists a no-shorting admissible portfolio $\pi_0(\cdot)$ such that

$$Z^*(t) = \sigma(t)'\pi_0(t), \text{ a.s., a.e.} t \in [0, T]. \tag{4.46}$$

Suppose (4.46) does not hold. Define

$$\pi_0(t) := \text{argmin}_{\pi \in \text{argmin}_{\pi \in \mathbf{R}^m_+} |\sigma(t)'\pi - Z^*(t)|^2} |\pi|^2.$$

Again, thanks to Lemma A.2, $\pi_0(\cdot)$ is a progressively measurable stochastic processes with respect to $\mathcal{F}_t$. Note $Z^*(t) \notin \{\sigma(t)'\pi : \pi \in \mathbf{R}^m_+\}$ whenever $\sigma(t)'\pi_0(t) - Z^*(t) \neq 0$.

Thus by Lemma B.3 and A.1, there is $\bar{D}(\cdot)$ which is progressively measurable satisfying $\bar{D}(t) \neq 0, Z^*(t)'\bar{D}(t) < 0, \sigma(t)\bar{D}(t) \geq 0$, a.s., a.e.$t$ on the set where $\sigma(t)'\pi_0(t) - Z^*(t) \neq 0$. Set

$$\hat{D}(t) := \begin{cases} 0, & \text{if } \sigma(t)'\pi_0(t) - Z^*(t) = 0, \\ \bar{D}(t)/|\bar{D}(t)|, & \text{if } \sigma(t)'\pi_0(t) - Z^*(t) \neq 0. \end{cases}$$

Then $\hat{D}(\cdot) \in L^\infty_{\mathcal{F}}(0, T, \mathbf{R}^n)$, $\sigma(t)\hat{D}(t) \geq 0$ and $E\int_0^T |\hat{D}(s)|^2 ds > 0$.

Set $N < K + 1$ with $|\theta^*(t)| \leq N,\ a.s.,\ a.e.t \in [0, T]$, $k = (K + 1 - N)/2$, $D = k\hat{D}$. Then $D(\cdot) \in L^\infty_{\mathcal{F}}(0, T, \mathbf{R}^n)$ with $\sigma(t)D(t) \geq 0$. Moreover,

$$Z^*(t)'D(t) < 0 \quad \text{whenever} \quad \sigma(t)'\pi_0(t) - Z^*(t) \neq 0. \tag{4.47}$$

Since $\sigma(t)[\theta^*(t) + D(t)] \geq \sigma(t)\theta^*(t) \geq B(t)$, and

$$|\theta^*(t) + D(t)| \leq |\theta^*(s)| + |D(s)| \leq N + (K + 1 - N)/2,\ a.s.,\ a.e.t \in [0, T],$$

we conclude $\theta^* + D \in \hat{\Theta}$.

Define $\bar{X}(\cdot)$ to be the solution of the following SDE:

$$\begin{cases} d\bar{X}(t) = [r(t)\bar{X}(t) + (\theta^*(t) + D(t))'Z^*(t)]dt + Z^*(t)'dW(t), \\ \bar{X}(0) = x_0. \end{cases}$$

Then

$$E[\bar{X}(T)H_{\theta^*+D}(T)] = \bar{X}(0) = x_0 = E[X^*(T)H_{\theta^*}(T)] \geq E[X^*(T)H_{\theta^*+D}(T)]. \tag{4.48}$$

On the other hand,

$$d[\bar{X}(t) - X^*(t)] = r(t)[\bar{X}(t) - X^*(t)]dt + Z^*(t)'D(t)dt,\ \bar{X}(0) - X^*(0) = 0;$$

hence $\bar{X}(T) - X^*(T) = \int_0^T e^{\int_t^T r(s)ds} Z^*(t)'D(t)dt$. It then follows from (4.48) and (4.47) that $\sigma(t)'\pi_0(t) - Z^*(t) = 0$, a.s., a.e.$t \in [0, T]$. This proves (4.46).

Next, let $\hat{X}(\cdot)$ be the solution to the following SDE:

$$\begin{cases} d\hat{X}(t) = [r(t)\hat{X}(t) + \theta_0(t)'Z^*(t)]dt + Z^*(t)'dW(t), \\ \hat{X}(t) = x_0. \end{cases} \tag{4.49}$$

Then $E[\hat{X}(T)H_{\theta_0}(T)] = x_0 \geq E[X^*(T)H_{\theta_0}(T)]$. On the other hand,

$$\begin{aligned} d[\hat{X}(t) - X^*(t)] = & \quad r(t)[\hat{X}(t) - X^*(t)]dt + [\theta_0(t) - \theta^*(t)]'Z^*(t)dt \\ = & \quad r(t)[\hat{X}(t) - X^*(t)]dt + [B(t) - \sigma(t)\theta^*(t)]'\pi_0(t)dt, \end{aligned}$$

where we have used the fact that $Z^*(t) = \sigma(t)'\pi_0(t)$. Hence,

$$\hat{X}(T) - X^*(T) = \int_0^T e^{\int_t^T r(s)ds}[B(t) - \sigma(t)\theta^*(t)]'\pi_0(t)dt \leq 0.$$

By $E[(\hat{X}(T) - X^*(T))H_{\theta_0}(T)] \geq 0$ we have

$$B(t)'\pi_0(t) = \theta^*(t)'\sigma(t)'\pi_0(t) \equiv \theta^*(t)'Z^*(t), \tag{4.50}$$

and

$$E[XH_{\theta^*}(T)] \equiv x_0 \equiv E[\hat{X}(T)H_{\theta_0}(T)] = E[XH_{\theta_0}(T)]. \tag{4.51}$$

It follows from (4.45) and (4.50) that $(X^*(\cdot), \pi_0(\cdot))$ satisfies

$$\begin{cases} dX^*(t) = [r(t)X^*(t) + B(t)'\pi_0(t)]dt + \pi_0(t)'\sigma(t)dW(t), \\ X^*(T) = X, \end{cases}$$

meaning that $X \in A_C$. Finally, the second assertion of the theorem follows from (4.51). $\square$

In view of Theorem 4.6.1, we can rewrite problem (4.8) for Case 4 as follows:

$$\begin{array}{ll} \text{Minimize} & E[X - (\lambda - \mu H_{\theta_0}(T))]^2, \\ \text{subject to} & \begin{cases} X \in L^2(\mathcal{F}_T, \mathbf{R}), \\ \sup_{\theta \in \hat{\Theta}_1} E[X(H_\theta(T) - H_{\theta_0}(T))] \leq 0, \\ X \geq 0. \end{cases} \end{array} \tag{4.52}$$

Denote $\hat{M} := \overline{\{H_\theta(T) - H_{\theta_0}(T) : \theta \in \hat{\Theta}_1\}}$, It is clear that $\hat{M} \subset L^2(\mathcal{F}_T, \mathbf{R})$, and $\hat{M}$ is a bounded convex set in view of Lemma 4.2.1.

Lemma 4.6.1 *For any $0 \leq k \leq 1$, if $Y \in \bar{M}$, then $kY \in \bar{M}$.*

Proof: For any $Y \in M$, denote $Y = H_\theta(T) - H_{\theta_0}(T)$. Then

$$\begin{aligned} kY &= kH_\theta(T) - kH_{\theta_0}(T) \\ &= kH_\theta(T) + (1-k)H_{\theta_0}(T) - H_{\theta_0}(T) \\ &= H_{\theta'}(T) - H_{\theta_0}(T) \\ &\in M. \end{aligned}$$

For any $Y \in \bar{M}$, there exists a sequence $Y_n \in M$ such that $Y_n \to Y$. Therefore $kY_n \in M$ and $kY_n \to kY$, which implies $kY \in \bar{M}$. $\square$

It is easy to see that problem(4.52) is equivalent to

$$
\begin{array}{ll}
\text{Minimize} & E[X-(\lambda-\mu H_{\theta_0}(T))]^2, \\
\text{subject to} & \begin{cases} X \in L^2(\mathcal{F}_T, \mathbf{R}), \\ \sup_{Y\in \bar{M}} E[XY] \le 0, \\ X \ge 0, \end{cases}
\end{array} \tag{4.53}
$$

which admits a unique optimal solution. For written convenience, denote $\hat{\lambda} = \lambda - \mu H_{\theta_0}(T)$.

Theorem 4.6.2 *Problem*

$$
\min_{Y\in \text{cone}\{\bar{M}\}} E[(\hat{\lambda}-Y)^+]^2 \tag{4.54}
$$

admits optimal solutions in $L^2(\mathcal{F}_T, \mathbf{R})$, and for different optimal solutions (if exist), $(\hat{\lambda} - Y^*)^+$ is the same, which is the unique optimal solution for problem (4.53) and problem (4.52).

Proof: The proof is the same as that of Theorem 4.5.3. □

Theorem 4.6.3 *Y^* is optimal for problem (4.54) if and only if $X^* = (\hat{\lambda} - Y^*)^+$ is no-shorting attainable and $E[X^*Y^*] = 0$.*

Proof: The proof is the same as that of part (i) in Theorem 4.4.2. □

We can regard problem (4.54) as the dual problem of problem (4.54) in Case 4. The static problem (4.7) in Case 4 is more difficult than in Case 2 and Case 3. But in the deterministic parameters case, we can still find the optimal solution explicitly by virtual of Theorem 4.6.3.

Let $\theta_1^*(\cdot)$ be the one defined in Section 4.4. It is easy to see that $\theta_1^* \in \hat{\Theta}_1$.

Proposition 4.6.1 *For any $\lambda \in \mathbf{R}, \mu \in \mathbf{R}^+$, $Y^* = \mu(H_{\theta_1^*}(T) - H_{\theta_0}(T))$ is optimal for problem (4.54). Therefore $X^* = [\lambda - \mu H_{\theta_1^*}(T)]^+$ is the unique optimal solution for problem (4.52).*

Proof: For any $\theta \in \hat{\Theta}$, $\theta(t)'\theta_1^*(t) \le |\theta_1^*(t)|^2$, $a.e.t \in [0,T]$. Define $\hat{W}(t) = W(t) + \int_0^t \theta_1^*(s)ds$, and let $\hat{E}$ be the expectation under the probability measure $\hat{P}$ defined by

$\frac{d\hat{P}}{dP} = e^{-\int_0^T |\theta_1^*(s)|^2/2ds - \int_0^T \theta_1^*(s)'dW(s)}$. Then by Girsanov's theorem, $\hat{W}(\cdot)$ is a standard Brownian motion. Therefore,

$$\begin{aligned}
& E[(\lambda - \mu H_{\theta_1^*}(T))^+ H_\theta(T)] \\
= \; & e^{-\int_0^T rds}\hat{E}(\lambda - \mu e^{-\int_0^T rds} e^{-\int_0^T |\theta_1^*|^2/2ds - \int_0^T \theta_1^*(s)'dW(s)})^+ \\
= \; & e^{-\int_0^T rds}\hat{E}(\lambda - \mu e^{-\int_0^T rds} e^{\int_0^T \theta'\theta_1^* ds} e^{-\int_0^T |\theta_1^*|^2/2ds - \int_0^T \theta_1^*(s)'d\hat{W}(s)})^+ \\
= \; & e^{-\int_0^T rds} E(\lambda - \mu e^{-\int_0^T rds} e^{\int_0^T \theta'\theta_1^* ds} e^{-\int_0^T |\theta_1^*|^2/2ds - \int_0^T \theta_1^*(s)'dW(s)})^+ \\
\leq \; & e^{-\int_0^T rds} E(\lambda - \mu e^{-\int_0^T rds} e^{\int_0^T |\theta_1^*|^2 ds} e^{-\int_0^T |\theta_1^*|^2/2ds - \int_0^T \theta_1^*(s)'dW(s)})^+ \\
= \; & E[(\lambda - \mu H_{\theta_1^*}(T))^+ H_{\theta_1^*}(T)],
\end{aligned}$$

and the equalities hold if and only if $\theta(\cdot)'\theta_1^*(\cdot) = |\theta_1^*(\cdot)|^2$. Call $\theta_1^*(t)'\theta_0(t) = |\theta_1^*(t)|^2$, therefore $E[X^*Y^*] = 0$. So we have $E[(\lambda - \mu H_{\theta_1^*}(T))^+(H_\theta(T) - H_{\theta_0})] \leq 0$. By Theorem 4.6.1, X^* is no-shorting attainable. The conclusion then follows Theorem 4.6.3. □

Theorem 4.6.4 *When $ze^{-\int_0^T rdt} > x_0 > 0$ and $\int_0^T \sum_{i=1}^n B(s)_i^+ ds > 0$, the optimal solution for problem (4.7) is $X^* = [\lambda - \mu H_{\theta_1^*}(T)]^+$, where (λ, μ) satisfies*

$$\begin{cases}
E[\lambda - \mu H_{\theta_1^*}(T)]^+ = z, \\
E[\lambda - \mu e^{\int_0^T |\theta_1^*|^2 dt} H_{\theta_1^*}(T)]^+ = x_0 e^{\int_0^T r(t)dt}.
\end{cases}$$

Moreover the efficient frontier is the following

$$\begin{cases}
E[x^*(T)] & = \dfrac{\eta e^{\int_0^T r(t)dt} N_1(\eta) - N_2(\eta)}{\eta N_2(\eta) - e^{-\int_0^T [r(s) - |\theta_1^*(s)|^2]ds} N_3(\eta)} x_0, \\
\text{Var } x^*(T) & = \left[\dfrac{\eta}{\eta N_1(\eta) - e^{-\int_0^T r(t)dt} N_2(\eta)} - 1\right][Ex^*(T)]^2 \\
& \quad - \dfrac{x_0}{\eta N_1(\eta) - e^{-\int_0^T r(t)dt} N_2(\eta)} Ex^*(T), \quad \eta \in (0, \infty],
\end{cases} \tag{4.55}$$

where

$$\begin{aligned}
N_1(\eta) &:= N\left(\frac{\ln \eta + \int_0^T [r(s) + \frac{1}{2}|\theta_1^*(s)|^2]ds}{\sqrt{\int_0^T |\theta_1^*(s)|^2 ds}}\right), \\
N_2(\eta) &:= N\left(\frac{\ln \eta + \int_0^T [r(s) - \frac{1}{2}|\theta_1^*(s)|^2]ds}{\sqrt{\int_0^T |\theta_1^*(s)|^2 ds}}\right), \\
N_3(\eta) &:= N\left(\frac{\ln \eta + \int_0^T [r(s) - \frac{3}{2}|\theta_1^*(s)|^2]ds}{\sqrt{\int_0^T |\theta_1^*(s)|^2 ds}}\right).
\end{aligned} \tag{4.56}$$

Proof: Using a similar argument in Section 3.3, we can prove when $z > 0, x_0 > 0$, the equations

$$\begin{cases}
E[\lambda - \mu H_{\theta_1^*}(T)]^+ = z, \\
E[\lambda - \mu H_{\theta_1^*}(T)]^+ H_{\theta_1^*}(T) = x_0
\end{cases}$$

admit a unique solution.

Comparing these two equations, along with the condition $ze^{\int_0^T r(s)ds} > x_0$, we find $\mu > 0$. By Proposition 4.6.3, $E[X^* H_{\theta_0}(T)] = E[X^* H_{\theta_1^*}(T)] = x_0$. By Theorems 4.6.1, 4.6.2 and 4.6.3, we conclude $X^* \in A_C$. By Theorems 4.2.2 and Theorem 4.6.2, X^* is the optimal solution for Problem (4.7).

The calculation of efficient frontier is the same as that in Sections 3.6 and 3.7. □

4.7 Conclusion

In this chapter, we have tried the decomposition method developed in Chapter 3 to study the mean-variance portfolio selection problem in an incomplete market. We studied two types of constraint, shorting prohibition and bankruptcy prohibition, and discuss four models generated by these two constraints.

From Section 4.3 to Section 4.6, we showed that the decomposition method worked, although more complicated, in an incomplete market. But the static optimization problem separated from the original problem is very difficult. In this chapter, we used the dual method to deal with the static optimization problem, and found the sufficient and necessary condition for a feasible solution to be optimal.

It should be noted that the dual problems of those static optimization problems are almost as difficult as their primal problems. But by the duality relation, we can get more information about their optimal solutions. At least the optimal solutions of the static problems can all be represented by the optimal solutions of their dual problems. When parameters of the market are deterministic, the optimal solutions for the dual problems can be found by the sufficient and necessary conditions, and then the static problems can be solved through their dual problems.

Lim [35] has studied the unconstrained mean-variance portfolio selection problem in an incomplete market in his recent work, where he applied the constrained linear-quadratic optimal control theory and BSDE to study the problem. The optimal solution for the unconstrained mean-variance problem in an incomplete market with stochastic parameters can be represented by the solution for a stochastic Riccati equation, whose existence of solution is proved. For the problem with no-bankruptcy constraint, Lim's method does not seem to work because the optimal control problem is one with a state constraint which is extremely difficult.

Chapter 5

General Mean-Risk Criteria

5.1 Introduction

Risk is a central issue in financial investment, yet it is a subjective notion as opposed to return. The practice of decision making under uncertainty frequently resorts to mean-risk model. Mean-variance is a classical mean-risk model. The mean-risk model allows a trade-off analysis between the mean level, which is sought, and the risk, which is aversive, of decision outcome.

A fundamental problem for mean-risk model is how risk should be measured. Risk measurement is a widely studied topic, which has been explored for a long time. In the middle of eighteenth century, Daniel Bernoulli already noted that risk should be a key fact for investment. In the early 1950s, Markowitz [41] proposed the single-period mean–variance (M–V) portfolio selection model, where he used the variance to measure the risk. This seminal work has been widely recognized to have laid the foundation of modern portfolio theory. However, there has also been substantial amount of objection to the measurement of risk by variance. The main aspects of the M–V theory under criticism include the penalty on the upside return, and the equal weight on the upside and downside whereas the asset return distribution is generally asymmetric. A critical argument against the mean-variance model is that it cannot be always consistent with stochastic dominance. See Ogryczak and Ruszczynski [50]. Consequently, some alternative risk measures were proposed, noticeably the so-called downside risk where only the return below its mean or a target level is counted as risk [16, 62, 49, 4]. One of the downside risk measures is the semivariance. In [42] Markowitz himself agreed that "semivariance seems more plausible than variance as a measure of risk". On the other hand, in a single-period financial market, other risk measures have also been proposed and studied, including VaR [22], mean–absolute deviation [28], and minimax measure [7]. For a recent survey on the Markowitz model and models with various risk measures, refer to [63].

The M–V approach "has received comparably little attention in the context of long-term investment planning" ([63, p.32]), especially in continuous time setting, until very recently. In a series of papers [70, 36, 34, 69] the continuous-time Markowitz models have been investigated thoroughly with closed-form solutions obtained in most cases. In this chapter, we will study continuous-time portfolio selection models with risk measures different from the variance. We will start with a weighted mean–variance problem where the risk has different weights on upside and downside returns. Explicit solution will be obtained for this model. While the weighted mean–variance model is important in its own right, it also converges to the mean–semivariance model when the weight on the upside variance goes to zero. Surprisingly and in sharp contrast to the single-period setting, based on this convergence approach we will show that the mean–semivariance model has *no* optimal solution, although asymptotically optimal solution can be obtained from the solution to the weighted mean–variance model. This "negative" result motivates us to study a general mean–downside-risk model where only the downside return is penalized, not necessarily in the fashion of variance. It turns out that this general downside-risk model provides no optimal solution either, under a very mild condition. It should be noted that Berlelarr [4] studied the downside-risk portfolio selection in a utility framework, and obtained optimal solutions. However, the results are obtained under an explicit constraint on the upper bound of the wealth. This constraint is not reasonable in the mean-risk framework.

Finally, we will study a "most general" mean–risk model, where the risk is measured by the expectation of a convex function of the deviation of the terminal payoff from its mean. For this model, we give a complete solution in terms of characterization of the existence of optimal portfolio and presentation of the solution when it exists. Furthermore, when there exists no optimal solution, we explore an asymptotic way to approach the infimum of the problem and the corresponding asymptotic solution sequence.

This chapter is organized as follows. In Section 5.2, we specify the continuous-time financial market under consideration, and introduce the equivalent static optimization problem for a dynamic portfolio selection problem. In Section 5.3, we investigate the weighted mean–variance problem, and in Section 5.4, we treat the mean–semivariance model based on the results in section 3 and a convergence approach. Section 5.5 is devoted to the study on the mean–downside-risk problem. In Section 5.6, we turn to the general mean–risk model, and find the sufficient and necessary conditions for the problem to admit optimal solutions, as well as the optimal solutions or asymptotic optimal solution sequences. Several examples are presented to illustrate the general results obtained. By the similar method, we solve the fixed-target portfolio selection problem in section 5.7. In Section 5.8, we study the existence of the optimal solution for mean-semivariance problem in a single-period, the conclusions are strikingly different from those in continuous-time market. Finally, Section 5.9 presents some concluding remarks.

5.2 Problem formulation

In Chapter 4, we have shown how to apply the decomposition idea in an incomplete market. In this chapter, we go back to the complete market specified in Section 3.2. Here we repeat some important notations and assumptions for this market.

Assumption 5.2.1

$$\sigma(t)\sigma(t)' \geq \delta I_m, \quad \forall t \in [0,T], \ \text{ a.s.} \tag{5.1}$$

for some $\delta > 0$, where I_m is the $m \times m$ identity matrix.

Define $\rho(\cdot)$ as the deflator process by

$$\rho(t) := \exp\left\{-\int_0^t [r(s) + \frac{1}{2}|\theta(s)|^2]ds - \int_0^t \theta(s)dW(s)\right\}. \tag{5.2}$$

With this definition, we know for any admissible portfolio defined in Definition 2.2.1, its wealth process $x(\cdot)$ satisfies

$$x(t) = \rho(t)^{-1} E(\rho(T)X(T)|\mathcal{F}_t). \tag{5.3}$$

The various portfolio selection models we are going to consider in this chapter are all special cases of the following general problem

$$\begin{array}{ll} \text{Minimize} & Ef(x(T) - Ex(T)), \\ \text{subject to} & \begin{cases} \pi(\cdot) \in L^2_{\mathcal{F}}(0, T; \mathbf{R}^m), \\ (x(\cdot), \pi(\cdot)) \text{ satisfies equation (2.5) with initial wealth } x_0, \\ Ex(T) = z, \end{cases} \end{array} \tag{5.4}$$

where $x_0, z \in \mathbf{R}$ and the function $f : \mathbf{R} \to \mathbf{R}$ are given. In words, problem (5.4) is to minimize the risk, measured by certain function of the deviation of the terminal wealth from its mean, via continuous trading, subject to an initial budget constraint (specified by x_0) and a target expected terminal payoff (specified by z). The Markowitz mean–variance problem is a special case of (5.4) with $f(x) = x^2$.

In model (5.4), we measure the risk of the terminal return of an investment by the expectation of a function f. The function f is supposed to be convex in general. The convexity of the function f is important for the existence of the optimal portfolio for model (5.4). With the convexity of f, an investor with the objective given in model (5.4) will, as suggested in the maxim, "not put all the eggs in one basket".

Thanks to the completeness of the market, we can apply the same technique used in Chapter 2 to decompose problem (5.4) into a static optimization problem and a wealth replication problem. The static optimization problem is

$$\begin{array}{ll} \text{Minimize} & Ef(X - z), \\ \text{subject to} & \begin{cases} EX = z, \\ E[\rho(T)X] = x_0, \\ X \in L^2(\mathcal{F}_T, \mathbf{R}). \end{cases} \end{array} \tag{5.5}$$

Suppose X^* is an optimal solution to (5.5), then the replication problem is to find a portfolio such that its terminal wealth hits X^*; in other words, the problem is to find

$(x(\cdot), \pi(\cdot))$ that solves the following equation

$$\begin{cases} dx(t) = [r(t)x(t)dt + B(t)\pi(t)]dt + \pi(t)'\sigma(t)dW(t), \\ x(T) = X^*. \end{cases} \tag{5.6}$$

Theorem 5.2.1 *If $(x^*(\cdot), \pi^*(\cdot))$ is optimal for problem (5.4), then $x^*(T)$ is optimal for problem (5.5) and $(x^*(\cdot), \pi^*(\cdot))$ satisfies (5.6). Conversely, if X^* is optimal for problem (5.5), then (5.6) must have a solution $(x^*(\cdot), \pi^*(\cdot))$ which is an optimal solution for (5.4).*

Proof: The proof is the same as that of Theorem 3.2.1. □

Remark 5.2.1 The replication problem (5.6) is essentially a backward stochastic differential equation; refer to [39, 40, 67] for various approaches in solving BSDEs. Indeed, the unique solution $(x^*(\cdot), \pi^*(t))$ of (5.6) is given by

$$\pi^*(t) = (\sigma(t)')^{-1}y^*(t), \tag{5.7}$$

whereas $(x^*(\cdot), y^*(\cdot))$ is the unique solution to the BSDE

$$\begin{cases} dx(t) = [r(t)x(t) + \theta(t)y(t)]dt + y(t)'dW(t), \\ x(T) = X^*. \end{cases} \tag{5.8}$$

Thus, according to Theorem 5.2.1 the key is to solve the static optimization problem (5.5). The remainder of this chapter will be mainly devoted to solving problem (5.5) for various situations.

5.3 The weighted mean–variance model

The classical mean–variance portfolio selection problem uses the variance as the measure for risk, which puts the same weight on the downside and upside (in relation to the mean) of the return. In this section, we study the "weighted" mean–variance portfolio selection model where the weights on the downside and upside may be different. Specifically, for given $\alpha > 0, \beta > 0, z \in \mathbf{R}, x_0 \in \mathbf{R}$, the problem being considered is

$$\begin{array}{ll} \text{Minimize} & J_1(x_0, z; \pi(\cdot)) := E[\alpha((x(T)-z)^+)^2 + \beta((x(T)-z)^-)^2], \\ \text{subject to} & \begin{cases} \pi(\cdot) \in L^2_{\mathcal{F}}(0,T;\mathbf{R}^m), \\ (x(\cdot), \pi(\cdot)) \text{ satisfies equation (2.5) with initial wealth } x_0, \\ Ex(T) = z. \end{cases} \end{array} \tag{5.9}$$

This model specializes (5.4) with $f(x) = \alpha(x^+)^2 + \beta(x^-)^2$. It reduces to the classical mean–variance model when $\alpha = \beta$.

As discussed at the end of Section 5.2, to solve the above problem it suffices to solve a static optimization problem (5.5) in terms of X. Define $Y := X - z$, then (5.5) specializes to

$$\begin{array}{ll} \text{Minimize} & E(\alpha(Y^+)^2 + \beta(Y^-)^2), \\ \text{subject to} & \begin{cases} EY = 0, \\ E[\rho Y] = y_0, \\ Y \in L^2(\mathcal{F}_T, \mathbf{R}), \end{cases} \end{array} \tag{5.10}$$

where $\rho := \rho(T)$ and $y_0 := x_0 - zE\rho$.

The above is a static convex optimization problem. Using the Lagrange multiplier approach (Proposition 3.4.1), we conclude that Y^* is an optimal solution of (5.10) if and only if Y^* is a feasible solution of (5.10) and there exists a pair (λ, μ) such that Y^* is an optimal solution of the following problem

$$\min_{Y \in L^2(\mathcal{F}_T, \mathbf{R})} E[\alpha(Y^+)^2 + \beta(Y^-)^2 - 2(\lambda - \mu\rho)Y]. \tag{5.11}$$

Lemma 5.3.1 *Problem (5.11) admits a unique optimal solution* $Y^* = \frac{(\lambda-\mu\rho)^+}{\alpha} - \frac{(\lambda-\mu\rho)^-}{\beta}$.

Proof: For any $Y \in L^2(\mathcal{F}_T, \mathbf{R})$, we have, samplewisely,

$$\begin{aligned} & \alpha(Y^+)^2 + \beta(Y^-)^2 - 2(\lambda - \mu\rho)Y \\ = \;& \alpha((Y^+)^2 - 2\frac{\lambda - \mu\rho}{\alpha}Y^+) + \beta((Y^-)^2 + 2\frac{\lambda - \mu\rho}{\beta}Y^-) \\ = \;& \alpha(Y^+ - \frac{\lambda - \mu\rho}{\alpha})^2 - \frac{(\lambda - \mu\rho)^2}{\alpha} + \beta(Y^- + \frac{\lambda - \mu\rho}{\beta})^2 - \frac{(\lambda - \mu\rho)^2}{\beta} \\ \geq \;& -\frac{((\lambda - \mu\rho)^+)^2}{\alpha} - \frac{((\lambda - \mu\rho)^-)^2}{\beta} \\ = \;& \alpha((Y^*)^+)^2 + \beta((Y^*)^-)^2 - 2(\lambda - \mu\rho)Y^*. \end{aligned}$$

This shows that Y^* is an optimal solution. The uniqueness of the optimal solution follows from the strict convexity of the problem (5.11). □

Proposition 5.3.1 *For any y_0, there exists a unique pair (λ, μ) such that the optimal solution Y^* in Lemma 5.3.1 satisfies $EY^* = 0, E[\rho Y^*] = y_0$. Moreover, $\lambda < 0, \mu < 0$ if $y_0 > 0$, $\lambda > 0, \mu > 0$ if $y_0 < 0$, and $\lambda = \mu = 0$ if $y_0 = 0$.*

Proof: If $y_0 = 0$, then we simply take $\lambda = \mu = 0$ (in which case $Y^* = 0$).

If $y_0 < 0$, then consider the following equation in the deterministic unknown ζ:

$$E(\zeta - \rho)^+/\alpha = E(\zeta - \rho)^-/\beta. \tag{5.12}$$

It is easy to see, using the mean-value theorem of a continuous function, that it admits a unique positive solution, denoted by $\zeta > 0$. Set

$$a := E[\rho(\zeta - \rho)^+]/\alpha - E[\rho(\zeta - \rho)^-]/\beta.$$

Note

$$\begin{aligned} E[\rho(\zeta - \rho)^+]/\alpha &= E[\rho(\zeta - \rho)^+ \mathbf{1}_{\rho<\zeta}]/\alpha \\ &< \zeta E(\zeta - \rho)^+/\alpha \\ &= \zeta E(\zeta - \rho)^-/\beta \\ &< E[\rho(\zeta - \rho)^- \mathbf{1}_{\rho>\zeta}]/\beta \\ &= E[\rho(\zeta - \rho)^-]/\beta. \end{aligned}$$

Hence $a < 0$. Take $\mu := y_0/a > 0, \lambda := \zeta\mu > 0$. Then it is straightforward that (λ, μ) is the desired pair.

Finally, if $y_0 > 0$, then let $\xi > 0$ be the unique solution of equation

$$E(\xi - \rho)^-/\alpha = E(\xi - \rho)^+/\beta. \tag{5.13}$$

Denoting

$$b := E[\rho(\xi - \rho)^-]/\alpha - E[\rho(\xi - \rho)^+]/\beta,$$

an argument similar to above yields $b > 0$. Take $\mu := -y_0/b < 0, \lambda := \xi\mu < 0$. Then (λ, μ) is the desired pair.

For the uniqueness, it is not difficult to prove by discussing for the cases $\mu < 0$ and $\mu > 0$ respectively. $\square$

Theorem 5.3.1 *The unique optimal solution for problem (5.10) is*

$$Y^* = \frac{(\lambda - \mu\rho)^+}{\alpha} - \frac{(\lambda - \mu\rho)^-}{\beta},$$

where (λ, μ) *is the unique solution of the system of equations:*

$$\begin{cases} \frac{E(\lambda-\mu\rho)^+}{\alpha} - \frac{E(\lambda-\mu\rho)^-}{\beta} = 0, \\ \frac{E[\rho(\lambda-\mu\rho)^+]}{\alpha} - \frac{E[\rho(\lambda-\mu\rho)^-]}{\beta} = y_0. \end{cases} \tag{5.14}$$

Moreover, the minimum value of the problem (5.10) is

$$E[\alpha((Y^*)^+)^2 + \beta((Y^*)^-)^2] = -\mu y_0 \equiv \begin{cases} -\frac{y_0^2}{E[\rho(\zeta-\rho)^+]/\alpha - E[\rho(\zeta-\rho)^-]/\beta}, & \text{if } y_0 > 0, \\ 0, & \text{if } y_0 = 0, \\ \frac{y_0^2}{E[\rho(\xi-\rho)^-]/\alpha - E[\rho(\xi-\rho)^+]/\beta}, & \text{if } y_0 < 0, \end{cases} \tag{5.15}$$

where ζ *and* ξ *are the solutions to (5.12) and (5.13) respectively.*

Proof: The first part of the theorem is immediate from Lemma 5.3.1 and Proposition 5.3.1. To prove the second part, note that the case when $y_0 = 0$ is trivial; so we consider $y_0 \neq 0$. One has

$$\begin{aligned} -\frac{1}{\beta}E((\lambda-\mu\rho)^-)^2 &= \frac{1}{\beta}E[(\lambda-\mu\rho)_-(\lambda-\mu\rho)] \\ &= \lambda\frac{E(\lambda-\mu\rho)^-}{\beta} - \mu\frac{E[\rho(\lambda-\mu\rho)^-]}{\beta} \\ &= \lambda\frac{E(\lambda-\mu\rho)^+}{\alpha} - \mu\left\{\frac{E[\rho(\lambda-\mu\rho)^+]}{\alpha} - y_0\right\} \\ &= \frac{1}{\alpha}E[(\lambda-\mu\rho)^+(\lambda-\mu\rho)] + \mu y_0 \\ &= \frac{1}{\alpha}E((\lambda-\mu\rho)^+)^2 + \mu y_0, \end{aligned}$$

where we have utilized the equations (5.14). Consequently,

$$\begin{aligned} E[\alpha((Y^*)^+)^2 + \beta((Y^*)^-)^2] &= \frac{1}{\alpha}E((\lambda-\mu\rho)^+)^2 + \frac{1}{\beta}E((\lambda-\mu\rho)^-)^2 \\ &= -\mu y_0. \end{aligned}$$

By the proof of Theorem 5.3.1, we have, when $y_0 < 0$,

$$\mu = \frac{y_0}{a} = \frac{y_0}{E[\rho(\zeta-\rho)^+]/\alpha - E[\rho(\zeta-\rho)^-]/\beta};$$

and when $y_0 > 0$,

$$\mu = -\frac{y_0}{b} = -\frac{y_0}{E[\rho(\xi-\rho)^-]/\alpha - E[\rho(\xi-\rho)^+]/\beta}.$$

Combining these equations, we obtain the desired result. $\square$

The following theorem gives a complete solution to the weighted mean–variance portfolio selection problem (5.9).

Theorem 5.3.2 *The unique optimal portfolio for (5.9) corresponding to $z > 0$ is given by*

$$\pi^*(t) = (\sigma(t)')^{-1} y^*(t), \tag{5.16}$$

where $(x^*(\cdot), y^*(\cdot))$ is the unique solution to the BSDE

$$\begin{cases} dx(t) = [r(t)x(t) + \theta(t)y(t)]dt + y(t)'dW(t) \\ x(T) = \frac{(\lambda-\mu\rho)^+}{\alpha} - \frac{(\lambda-\mu\rho)^-}{\beta} + z, \end{cases} \tag{5.17}$$

with (λ, μ) being the unique solution to the system of algebraic equations

$$\begin{cases} \frac{E(\lambda-\mu\rho)^+}{\alpha} - \frac{E(\lambda-\mu\rho)^-}{\beta} = 0 \\ \frac{E[\rho(\lambda-\mu\rho)^+]}{\alpha} - \frac{E[\rho(\lambda-\mu\rho)^-]}{\beta} = x_0 - zE\rho. \end{cases} \tag{5.18}$$

Moreover, the minimum value of (5.9), as a function of $(x_0, z) \in \mathbf{R} \times \mathbf{R}$, is given by

$$J_1^*(x_0, z) = \begin{cases} -\frac{(x_0-zE\rho)^2}{E[\rho(\zeta-\rho)^+]/\alpha - E[\rho(\zeta-\rho)^-]/\beta}, & \textit{if } x_0 - zE\rho > 0, \\ 0, & \textit{if } x_0 - zE\rho = 0, \\ \frac{(x_0-zE\rho)^2}{E[\rho(\xi-\rho)^-]/\alpha - E[\rho(\xi-\rho)^+]/\beta}, & \textit{if } x_0 - zE\rho < 0, \end{cases} \tag{5.19}$$

where ζ and ξ are the solutions to (5.12) and (5.13) respectively.

Proof: It has been proved that the optimal solution to the static optimization problem (5.5) in the present case is $X^* = \frac{(\lambda-\mu\rho)^+}{\alpha} - \frac{(\lambda-\mu\rho)^-}{\beta} + z$. On the other hand, (5.17) is a linear BSDE with uniformly bounded linear coefficients; hence it admits a unique solution $(x^*(\cdot), y^*(\cdot))$. As a result, $(x^*(\cdot), \pi^*(\cdot)) \equiv (x^*(\cdot), (\sigma(\cdot)')^{-1}y^*(\cdot))$ solves the replication problem (5.6). Finally, (5.19) follows readily from (5.15). $\square$

Remark 5.3.1 If $z = \frac{x_0}{E\rho}$, then $\lambda = \mu = 0$ implying that $x^*(T) = z$ a.s. under the optimal portfolio. Hence in this case the optimal portfolio is a risk-free portfolio. As a by-product, we have proved that a risk-free portfolio is available (which involves exposure to the stocks) even though the interest rate is random.

Remark 5.3.2 When the market coefficients are deterministic, optimal portfolio can be obtained more explicitly via some Black–Scholes type equation as in Section 3.7.

5.4 The mean–semivariance model

In this section we consider the mean–semivariance problem, where only the downside return is penalized. For each $z \in \mathbf{R}$, the problem is to

$$\begin{array}{ll} \text{Minimize} & E[((x(T)-z)^-)^2], \\ \text{subject to} & \begin{cases} \pi(\cdot) \in L^2_{\mathcal{F}}(0,T;\mathbf{R}^m), \\ (x(\cdot),\pi(\cdot)) \text{ satisfies equation (2.5) with initial wealth } x_0, \\ Ex(T) = z. \end{cases} \end{array} \tag{5.20}$$

This is a case of (5.4) with $f(x) = (x^-)^2$.

Denote $\rho := \rho(T)$ where $\rho(\cdot)$ is defined by (5.2). As in Section 5.3, we define

$$\rho_0 := \inf\{\eta \in \mathbf{R} : P(\rho < \eta) > 0\}, \quad \rho_1 := \sup\{\eta \in R : P(\rho > \eta) > 0\}. \tag{5.21}$$

Lemma 5.4.1 *Let $\zeta(\alpha)$, $\alpha \in (0,1)$, be the solution to (5.12) with $\beta = 1-\alpha$, then* $\lim_{\alpha\downarrow 0}\zeta(\alpha) = \rho_0$. *Similarly, let $\xi(\alpha)$, $\alpha \in (0,1)$, be the solution to (5.13) with $\beta = 1-\alpha$, then* $\lim_{\alpha\downarrow 0}\xi(\alpha) = \rho_1$.

Proof: Define $f(\zeta) := \frac{E(\zeta-\rho)^+}{E(\zeta-\rho)^-}$, $\zeta \in (\rho_0,\rho_1)$. Then equation (5.12) is equivalent to $f(\zeta) = \frac{\alpha}{1-\alpha}$. Obviously, $f(\zeta)$ is a strictly positive and strictly increasing function on $\zeta \in (\rho_0,\rho_1)$; hence $\zeta(\alpha)$ is strictly increasing on $\alpha \in (0,1)$, and in this interval, $\rho_0 < \zeta(\alpha) < \rho_1$.

Denote $\lim_{\alpha\downarrow 0}\zeta(\alpha) = \zeta_0$, then $\zeta_0 \geq \rho_0$. If $\zeta_0 > \rho_0$, then take $\zeta \in (\rho_0,\zeta_0)$. Since $\zeta < \zeta_0 = \lim_{\alpha\downarrow 0}\zeta(\alpha)$, we have $\frac{E(\zeta-\rho)^+}{E(\zeta-\rho)^-} = 0$, implying $E(\zeta-\rho)^+ = 0$. However, $\zeta > \rho_0$, so $P(\rho < \zeta) > 0$ leading to a contradiction. Therefore, $\zeta_0 = \rho_0$.

Similarly, we can prove the other part of the lemma in terms of $\xi(\alpha)$. □

We are now in a position to prove the following *negative* result.

Theorem 5.4.1 *The mean–semivariance problem (5.20) does not admit any optimal solution so long as $z \neq \frac{x_0}{E\rho}$.*

Proof: In view of Theorem 5.2.1, it suffices to prove that the static optimization problem

$$\begin{array}{ll} \text{Minimize} & E((Y^-)^2), \\ \text{subject to} & \begin{cases} EY = 0, \\ E[\rho Y] = y_0 \equiv x_0 - zE\rho, \\ Y \in L^2(\mathcal{F}_T, \mathbf{R}), \end{cases} \end{array} \tag{5.22}$$

has no optimal solution. Consider problem (5.10) with $\beta = 1 - \alpha$ and $\alpha \in (0, 1)$. It has been proved in the proof of Proposition 5.3.1 that there exists a pair $(\lambda(\alpha), \mu(\alpha))$ such that $Y(\alpha) = \frac{(\lambda(\alpha)-\mu(\alpha)\rho)^+}{\alpha} - \frac{(\lambda(\alpha)-\mu(\alpha)\rho)^-}{\beta}$ satisfies $EY(\alpha) = 0, E[\rho Y(\alpha)] = y_0$. This implies that each $Y(\alpha)$ is feasible for problem (5.22).

Since $z \neq \frac{x_0}{E\rho}$, we have $y_0 \neq 0$. First consider the case when $y_0 < 0$. It was proved in the proof of Proposition 5.3.1 that $\lambda(\alpha) > 0, \mu(\alpha) > 0$. Let $\zeta(\alpha) = \lambda(\alpha)/\mu(\alpha)$. Then $\zeta(\alpha)$ is the solution to (5.12) with $\beta = 1 - \alpha$. Lemma 5.4.1 along with its proof yields $\zeta(\alpha) > \rho_0$, and $\zeta(\alpha) \to \rho_0$ as $\alpha \downarrow 0$. However,

$$\begin{aligned} 0 \leq E[(\rho-\rho_0)(\zeta(\alpha)-\rho)^+]/\alpha &\leq (\zeta(\alpha)-\rho_0)E(\zeta(\alpha)-\rho)^+/\alpha \\ &= (\zeta(\alpha)-\rho_0)E(\zeta(\alpha)-\rho)^-/(1-\alpha) \\ &\leq (\zeta(\alpha)-\rho_0)E\rho/(1-\alpha) \to 0, \quad \text{as } \alpha \downarrow 0, \end{aligned}$$

and

$$E[(\rho-\rho_0)(\zeta(\alpha)-\rho)^-]/(1-\alpha) \to E(\rho-\rho_0)^2, \quad \text{as } \alpha \downarrow 0.$$

Consequently,

$$\mu(\alpha) \equiv \frac{y_0}{E[(\rho-\rho_0)(\zeta(\alpha)-\rho)^+]/\alpha - E[(\rho-\rho_0)(\zeta(\alpha)-\rho)^-]/(1-\alpha)} \to -y_0/E(\rho-\rho_0)^2, \quad \text{as } \alpha \downarrow 0,$$

and, therefore,

$$E[(Y(\alpha)^-)^2] = \frac{\mu(\alpha)^2 E((\zeta(\alpha)-\rho)^-)^2}{(1-\alpha)^2} \to y_0^2/E(\rho-\rho_0)^2, \quad \text{as } \alpha \downarrow 0. \tag{5.23}$$

On the other hand, for any feasible solution Y of problem (5.22), Cauchy–Schwartz's inequality yields $\{E[(\rho-\rho_0)Y^-]\}^2 \leq E[Y^-]^2 E[(\rho-\rho_0)^2 \mathbf{1}_{Y<0}]$. Note that $E[(\rho-\rho_0)^2\mathbf{1}_{Y<0}] \neq 0$, for otherwise $P(Y \geq 0) = 1$ which together with $EY = 0$ implies $P(Y = 0) = 1$ and hence $y_0 = 0$. As a result,

$$E[Y^-]^2 \geq \frac{\{E[(\rho-\rho_0)Y^-]\}^2}{E[(\rho-\rho_0)^2\mathbf{1}_{Y<0}]} = \frac{\{E[(\rho-\rho_0)Y^+] - y_0\}^2}{E[(\rho-\rho_0)^2\mathbf{1}_{Y<0}]} > \frac{y_0^2}{E(\rho-\rho_0)^2}, \tag{5.24}$$

where the last *strict* inequality is due to the facts that $y_0 < 0$ and $EY = 0$. Comparing (5.23) and (5.24) we conclude that there is no optimal solution for (5.22) in this case.

For the case $y_0 > 0$, we have proved that $\lambda(\alpha) < 0, \mu(\alpha) < 0$ and $\xi(\alpha) := \lambda(\alpha)/\mu(\alpha) > 0$, where $\xi(\alpha)$ is the solution to (5.13) with $\beta = 1-\alpha$. According to Lemma 5.4.1, $\xi(\alpha) \to \rho_1$

as $\alpha \downarrow 0$. First assume that $\rho_1 < +\infty$. Then an argument completely analogous to the above yields

$$E[(Y(\alpha)^-)^2] \to y_0^2/E(\rho_1-\rho)^2, \quad \text{as } \alpha \downarrow 0, \tag{5.25}$$

whereas $E[Y^-]^2 > y_0^2/E(\rho_1-\rho)^2$ for any feasible solution Y of problem (5.22). Thus there is no optimal solution for (5.22).

If $\xi(\alpha) \to \rho_1 = +\infty$, then

$$\begin{aligned} & E[\rho(\xi(\alpha)-\rho)^-]/\alpha - E[\rho(\xi(\alpha)-\rho)^+]/(1-\alpha) \\ \geq \;& \xi(\alpha)E(\xi(\alpha)-\rho)^-/\alpha - E[\rho(\xi(\alpha)-\rho)^+]/(1-\alpha) \\ = \;& \xi(\alpha)E(\xi(\alpha)-\rho)^+/(1-\alpha) - E[\rho(\xi(\alpha)-\rho)^+]/(1-\alpha) \\ = \;& E((\xi(\alpha)-\rho)^+)^2/(1-\alpha) \\ \to \;& +\infty, \quad \text{as } \alpha \downarrow 0. \end{aligned} \tag{5.26}$$

However,

$$\begin{aligned} E[(Y(\alpha)^-)^2] = \;& \tfrac{\mu(\alpha)^2 E((\xi(\alpha)-\rho)^+)^2}{(1-\alpha)^2} \\ \leq \;& \tfrac{\mu(\alpha)^2}{1-\alpha}\{E[\rho(\xi(\alpha)-\rho)^-]/\alpha - E[\rho(\xi(\alpha)-\rho)^+]/(1-\alpha)\} \\ = \;& \tfrac{y_0^2}{(1-\alpha)\{E[\rho(\xi(\alpha)-\rho)^-]/\alpha - E[\rho(\xi(\alpha)-\rho)^+]/(1-\alpha)\}} \\ \to \;& 0 \quad \text{as } \alpha \downarrow 0, \end{aligned} \tag{5.27}$$

where the first inequality was due to (5.26) and the second equality was because of the fact that $\mu(\alpha) = -\frac{y_0}{E[\rho(\xi(\alpha)-\rho)^-]/\alpha - E[\rho(\xi(\alpha)-\rho)^+]}$. On the other hand, for any feasible solution Y , if $E(Y^-)^2 = 0$, then $Y = 0$ implying $y_0 = 0$. This, once again, proves that (5.22) has no optimal solution. □

Remark 5.4.1 If $z = \frac{x_0}{E\rho}$, then there is a risk-free portfolio under which the terminal wealth is exactly z; see Remark 5.3.1. This portfolio is therefore an optimal portfolio for (5.22).

Remark 5.4.2 Although the mean–semivariance problem in general do not admit optimal solutions, the infimum of the problem has been obtained explicitly in the proof of Theorem 5.4.1. Specifically, the infimum is $\frac{y_0^2}{E(\rho-\rho_0)^2}$ if $y_0 < 0$, and is $\frac{y_0^2}{E(\rho_1-\rho)^2}$ if $y_0 > 0$. Moreover, asymptotically optimal portfolios can be obtained by replicating $Y(\alpha)$ as $\alpha \to 0$.

Theorem 5.4.1 shows that, quite contrary to the single-period case, the mean–semivariance portfolio selection problem in a complete continuous-time financial market does not admit a solution (save for the trivial case when $z = \frac{x_0}{E\rho}$). In the next section, we shall extend this "negative" result to a general model that concerns only the downside risk.

5.5 The mean–downside-risk model

Some alternative measures for risk have been proposed in lieu of the variance, and one of such measures is the downside risk which concerns only the downside deviation of the return from the mean. The semivariance studied in the previous section is a typical type of downside risk measure. In this section, we will generalize the result obtained in Section 5.4 to a general portfolio selection model with downside risk.

Before we formulate the underlying portfolio selection problem, let us investigate an abstract static optimization problem, which is interesting in its own right. Let $(\Omega, \mathcal{F}, P)$ be a probability space. For $q \geq 1$, we denote by $L^q(\mathcal{F}, \mathbf{R})$ the set of all $\mathcal{F}$-measurable real random variables X such that $|X|^q$ is integrable under P. Let ξ be a strictly positive real random variable, with the property that

$$P\{\xi \in (M_1, M_2)\} > 0, \;\text{ and }\; P\{\xi = M_1\} = P\{\xi = M_2\} = 0, \; \forall 0 \leq M_1 < M_2 \leq +\infty. \tag{5.28}$$

Consider the following optimization problem, with a given $y_0 \in \mathbf{R}$:

$$\begin{array}{ll} \text{Minimize} & Ef(Y), \\ \text{subject to} & \begin{cases} EY = 0, \\ E[\xi Y] = y_0, \\ Y \in L^q(\mathcal{F}, \mathbf{R}), \end{cases} \end{array} \tag{5.29}$$

where $f : \mathbf{R} \to \mathbf{R}$ is a given function. Throughout this section we impose the following assumption on f:

Assumption 5.5.1 $f \geq 0$, left continuous at 0, strictly decreasing on $\mathbf{R}^-$, and $f(x) = 0$ $\forall x \in \mathbf{R}^+$.

Notice that here we even do not need the convexity of f, which will be an essential requirement in the next section. An example of such a function is $f(x) = (x^-)^p$ for some $p \geq 0$. By virtue of the assumed properties of f, problem (5.29) has a finite (nonnegative, in fact) infimum.

Theorem 5.5.1 *Problem (5.29) admits no optimal solution for any $y_0 \neq 0$.*

This theorem will be proved via several intermediate results. Denote $L^q(\mathcal{F}, \mathbf{R}^-) := \{X \in L^q(\mathcal{F}, \mathbf{R}) : X \leq 0\}$. For any $a \leq 0$, define

$$h(\alpha) := \inf_{Z \in L^q(\mathcal{F}, \mathbf{R}^-),\ E[\xi Z] = \alpha} Ef(Z).$$

Lemma 5.5.1 *$h(\alpha)$ is decreasing on $\mathbf{R}^-$. Moreover, if for a given $\alpha_1 < 0$, there exists $\bar{Z} \in L^q(\mathcal{F}_T, \mathbf{R}^-)$ such that $E[\xi \bar{Z}] = \alpha_1, Ef(\bar{Z}) = h(\alpha_1)$, then $h(\alpha_1) > h(\alpha_2) \quad \forall \alpha_2 \in (\alpha_1, 0)$.*

Proof: For any $\alpha_1 < \alpha_2 < 0$, we have

$$h(\alpha_2) \leq \inf_{Z \in L^q(\mathcal{F}_T, \mathbf{R}^-),\ E[\xi Z] = \alpha_1} Ef(\frac{\alpha_2}{\alpha_1} Z) \leq \inf_{Z \in L^q(\mathcal{F}_T, \mathbf{R}^-),\ E[\xi Z] = \alpha_1} Ef(Z) = h(\alpha_1).$$

If there exists a $\bar{Z} \in L^q(\mathcal{F}_T, \mathbf{R}^-)$ with $E[\xi \bar{Z}] = \alpha_1, Ef(\bar{Z}) = h(\alpha_1)$, then

$$h(\alpha_2) \leq Ef(\frac{\alpha_2}{\alpha_1} \bar{Z}) < Ef(\bar{Z}) = h(\alpha_1).$$

This completes the proof. $\square$

Lemma 5.5.2 *For any $x > 0$, $\delta > 0$, and $0 < y < x\delta$, there exists a uniformly bounded random variable $\bar{Y} \geq 0$ such that $E\bar{Y} = x, E[\xi \bar{Y}] = y$, and $\bar{Y} = 0$ on the set $\{\omega \in \Omega : \xi \geq \delta\}$.*

Proof: Take $\delta_1 < \delta_2 < \delta$ so that $E(\xi | \delta_1 \leq \xi < \delta_2) = y/x$. The property of the distribution of ξ and the fact that $y/x < \delta$ ensure the existence of such δ_1, δ_2. Define $\bar{Y} = \frac{x}{P(\delta_1 \leq \xi < \delta_2)} \mathbf{1}_{\delta_1 \leq \xi < \delta_2}$. Then $\bar{Y}$ satisfies all the desired requirements. $\square$

Lemma 5.5.3 *For any $y_0 < 0$ and $\epsilon > 0$, there exists a feasible solution Y for problem (5.29) such that $Ef(Y) < h(y_0) + \epsilon$.*

Proof: For any $\epsilon > 0$, there exists $Z \in L^q(\mathcal{F}_T, \mathbf{R}^-)$ such that $E[\xi Z] = y_0$, and $h(y_0) \leq Ef(Z) < h(y_0) + \epsilon$. Since $\frac{\alpha}{y_0} E[\xi Z] = \alpha \ \forall \alpha < y_0$, we have $h(\alpha) \leq Ef(\frac{\alpha}{y_0} Z)$. Fix $\alpha < y_0$. Since the distribution of ξ has no atom by the assumption, there exists $\delta_0(\alpha) > 0$ such that

$$\frac{\alpha}{y_0} E[Z \xi \mathbf{1}_{\xi \geq \delta_0(\alpha)}] = y_0.$$

As a result, one can take $\delta_1(\alpha) > 0$ with $\delta_1(\alpha) < \delta_0(\alpha)$ and

$$\frac{-E[\frac{\alpha}{y_0}Z\mathbf{1}_{\xi\geq\delta_1(\alpha)}]}{y_0 - \frac{\alpha}{y_0}E[Z\xi\mathbf{1}_{\xi\geq\delta_1(\alpha)}]} > \frac{1}{\delta_1(\alpha)}.$$

It is easy to see that $\lim_{a\uparrow y_0}\delta_0(\alpha) = 0$; hence $\lim_{a\uparrow y_0}\delta_1(\alpha) = 0$.

Define

$$Y_\alpha = \begin{cases} \frac{\alpha}{y_0}Z, & \text{if } \xi \geq \delta_1(\alpha), \\ \bar{Y}_\alpha, & \text{if } \xi < \delta_1(\alpha), \end{cases}$$

where $\bar{Y}_\alpha \geq 0$ is such that $\bar{Y}_\alpha = 0$ on the set $\{\omega \in \Omega : \xi \geq \delta_1(\alpha)\}$, and

$$\begin{aligned} E\bar{Y}_\alpha &= -E[\tfrac{\alpha}{y_0}Z\mathbf{1}_{\xi\geq\delta_1(\alpha)}], \\ E[\xi\bar{Y}_\alpha] &= y_0 - E[\tfrac{\alpha}{y_0}\xi Z\mathbf{1}_{\xi\geq\delta_1(\alpha)}]. \end{aligned}$$

The existence of such $\bar{Y}_\alpha$ is implied by Lemma 5.5.2. Consequently, $EY_\alpha = 0, E[\xi Y_\alpha] = y_0$, meaning that Y_α is feasible for problem (5.29).

Now $Ef(Y_\alpha) = E[f(\frac{\alpha}{y_0}Z)\mathbf{1}_{\xi\geq\delta_1(\alpha)}] + E[f(\bar{Y}_\alpha)\mathbf{1}_{\xi<\delta_1(\alpha)}] = E[f(\frac{\alpha}{y_0}Z)\mathbf{1}_{\xi\geq\delta_1(\alpha)}]$. Thus, we have

$$Ef(\frac{\alpha}{y_0}Z) \geq Ef(Y_\alpha) \geq E[f(Z)\mathbf{1}_{\xi\geq\delta_1(\alpha)}]$$

which implies $\lim_{a\uparrow\eta} Ef(Y_\alpha) = Ef(Z) < h(y_0) + \epsilon$. Thus, we can take $a < y_0$ such that $Ef(Y_\alpha) < h(y_0) + \epsilon$. $\square$

Proposition 5.5.1 *Problem (5.29) admits no optimal solution for any $y_0 < 0$.*

Proof: In view of Lemma 5.5.3 it suffices to show that $Ef(Y) > h(y_0)$ for any feasible solution Y of (5.29). To this end, first note that $E[\xi Y^+] > 0$, for otherwise $Y^+ = 0$ which along with $EY = 0$ yields $Y = 0$ and hence $y_0 = 0$. Therefore, $\alpha := E[-\xi Y^-] < y_0$, suggesting $h(\alpha) \geq h(y_0)$ by virtue of Lemma 5.5.1. If $h(\alpha) = h(y_0)$, then the contrapositive of Lemma 5.5.1 implies that $Ef(-Y^-) > h(\alpha)$. So $Ef(Y) \geq Ef(-Y^-) > h(\alpha) = h(y_0)$. Otherwise, if $h(\alpha) > h(y_0)$, then $Ef(Y) \geq h(\alpha) > h(y_0)$. $\square$

Now let us turn to the case when $y_0 > 0$.

Proposition 5.5.2 *Problem (5.29) admits no optimal solution for any $y_0 > 0$.*

Proof: Since $y_0 > 0$, any feasible solution Y of problem (5.29) satisfies $Ef(Y) > 0$. Thus we only need to show that there exists a sequence $\{Y_n\}$ of feasible solutions for problem (5.29) with $\lim_{n\to+\infty} Ef(Y_n) = 0$. Indeed, for any $n > 0$, define

$$Y_n = \begin{cases} -\alpha_n, & \text{if } \xi < n, \\ \beta_n, & \text{if } \xi \geq n, \end{cases}$$

where α_n, β_n are defined by

$$\alpha_n = \frac{y_0}{[E(\xi|\xi \geq n) - E(\xi|\xi < n)]P(\xi < n)}, \quad \beta_n = \frac{y_0}{[E(\xi|\xi \geq n) - E(\xi|\xi < n)]P(\xi \geq n)}.$$

Then it is easy to verify that $\alpha_n > 0, \beta_n > 0$, $\lim_{n\to+\infty} \alpha_n = 0$, and $EY_n = 0, E[\xi Y_n] = y_0$. Thus, $\{Y_n\}$ are feasible solutions for (5.29), and

$$0 \leq Ef(Y_n) = E[f(-\alpha_n)\mathbf{1}_{\xi<n}] \leq f(-\alpha_n).$$

Since f is left continuous at 0, we conclude $\lim_{n\to+\infty} Ef(Y_n) = 0$. $\square$

Remark 5.5.1 In the proof of Proposition 5.5.2, only the following properties of $f(\cdot)$ was utilized: $f(x) > 0$ if $x < 0$, $f(x) = 0$ if $x \geq 0$, and $\lim_{x\uparrow 0} f(x) = 0$. The strictly decreasing property of $f(\cdot)$ was not necessary.

Combining Proposition 5.5.1 and Proposition 5.5.2. we obtain the conclusion of Theorem 5.5.1.

Now we turn to the continuous-time portfolio selection problem (5.4) where f satisfies Assumption 5.5.1. The way the function f is given suggests that only the downside deviation of the terminal wealth from its mean is penalized; hence the model constitutes a (very general) mean–downside-risk portfolio selection problem.

Let $\rho(\cdot)$ be the price kernel defined by (5.2). We impose the following assumption:

Assumption 5.5.2 For any $0 \leq M_1 < M_2 \leq +\infty$, $P\{\rho(T) \in (M_1, M_2)\} > 0$ and $P\{\rho(T) = M_1\} = P\{\rho(T) = M_2\} = 0$.

This assumption is satisfied when, say, $r(\cdot)$ and $\theta(\cdot)$ are deterministic and $\int_0^T |\theta(t)|^2 dt > 0$.

The corresponding static optimization problem (5.5), after taking a transformation $Y := X - z$, is exactly the problem (5.29) with $q = 2$. Hence, by virtue of Theorems 5.5.1 and 5.2.1, we conclude the following result.

Theorem 5.5.2 *Under Assumption (5.5.1) and 5.5.2, Problem (5.4) admits no optimal solution for any* $z \neq \frac{x_0}{E\rho(T)}$. *On the other hand, if* $z = \frac{x_0}{E\rho(T)}$, *then (5.4) has an optimal portfolio which is the risk-free portfolio.*

Theorem 5.5.2 claims that a mean–downside-risk portfolio selection problem is generally not well-posed in a complete continuous-time financial market. It is a very general result; however it does not completely cover Theorem 5.4.1 since the latter does not require Assumption 5.5.2.

5.6 General mean–risk model

We have shown in the last section that in the continuous-time setting, the mean–downside-risk model is in general not well-posed. In other words, problem (5.4) does not admit an optimal solution if the function f has the property that it vanishes on the nonnegative half real axis. Notice that for this negative result to hold the function f is *not* required to be convex. In this section, we will study model (5.4) where a general *convex* function f is used to measure the risk. We will give a complete solution to the problem in terms of telling exactly when the portfolio selection problem possesses an optimal solution and, when it does, giving the explicit solution.

5.6.1 A static optimization problem

As in the previous sections, we first need to investigate a static optimization problem. Let $(\Omega, \mathcal{F}, P)$ be a probability space and ξ a strictly positive real random variable on it satisfying (5.28). Consider a convex (hence continuous) function $f : \mathbf{R} \to \mathbf{R}$, not necessarily differentiable. For any $x \in \mathbf{R}$, its subdifferential $\partial f(x)$ in the sense of convex analysis (see, e.g., [56]), is defined as the set

$$\partial f(x) := \{x^* \in \mathbf{R} : \ f(y) - f(x) \geq x^*(y - x), \ \ \forall y \in \mathbf{R}\} \equiv [f'_-(x), f'_+(x)], \tag{5.30}$$

where $f'_-(x)$ and $f'_+(x)$ are the left and right derivatives of f at x respectively. The set $\partial f(x)$ is a nonempty bounded set for every $x \in \mathbf{R}$ ([56, Theorem 23.4]). Moreover, the convexity of f implies that the subdifferential is non-decreasing in the sense that

$$f'_+(x_1) \leq f'_-(x_2), \ \ \forall x_1 \leq x_2. \tag{5.31}$$

We call a convex function f to be strictly convex at $x_0 \in \mathbf{R}$ if

$$f(x_0) < \kappa f(x_1) + (1-\kappa) f(x_2)$$

for any $x_1 < x_0 < x_2$ and $\kappa \in (0,1)$ with $\kappa x_1 + (1-\kappa)x_2 = x_0$. A convex function is called strictly convex if it is strictly convex at every $x \in \mathbf{R}$. Some properties of a convex function that are useful in this chapter are presented in Appendix C.

For a given $q \geq 1$ and $y_0 \in \mathbf{R}$, consider the following optimization problem:

$$\begin{array}{ll} \text{Minimize} & Ef(Y), \\ \text{subject to} & \begin{cases} EY = 0, \\ E[\xi Y] = y_0, \\ Y \in L^q(\mathcal{F}, \mathbf{R}), \end{cases} \end{array} \tag{5.32}$$

where $f : \mathbf{R} \to \mathbf{R}$ is a given function. Throughout this section we assume that f satisfies

Assumption 5.6.1 f is convex, and strictly convex at 0.

Note that the strict convexity at 0 is a very mild condition, which is valid in many meaningful cases (see the examples in Section 5.6.4).

In view of Jensen's inequality, one has $Ef(Y) \geq f(EY) = f(0)$ for any feasible solution Y of (5.32). Hence problem (5.32) has a finite infimum if its feasible region is nonempty. Also we see that if $y_0 = 0$, then (5.32) has (trivially) an optimal solution $Y^* = 0$ a.s.. On the other hand, due to the convexity of f, we can apply Proposition 3.4.1 to conclude that (5.32) admits an optimal solution Y^* if and only if Y^* is feasible for (5.32) and there exists a pair (λ, μ) such that Y^* solves the following problem

$$\min_{Y \in L^q(\mathcal{F}, \mathbf{R})} E[f(Y) - (\lambda - \mu\xi)Y]. \tag{5.33}$$

Lemma 5.6.1 $Y^* \in L^q(\mathcal{F}, \mathbf{R})$ *is an optimal solution to (5.33) if and only if*

$$f(Y^*) - (\lambda - \mu\xi)Y^* = \min_{y \in \mathbf{R}} [f(y) - (\lambda - \mu\xi)y], \quad \text{a.s..}$$

Proof: The "if" part is obvious. We now prove the "only if" part. Suppose $Y^* \in L^q(\mathcal{F}, \mathbf{R})$ is an optimal solution to (5.33). Define $h(y) := f(y) - (\lambda - \mu\xi)y, \quad y \in \mathbf{R}$, and $c := \inf_{y \in \mathbf{R}} h(y)$. Let $Z := \cup_{n \in \mathbf{N}} \{(z_1, \cdots, z_n) : z_i \in \mathbf{Q}\}$, and $\bar{h}(z) := \inf_{1 \leq i \leq n} h(z_i, \omega)$ for

$z = (z_1, \cdots, z_n) \in Z$. Since $h(y)$ is continuous in y, we have $c = \inf_{z \in Z} \bar{h}(z)$. Now, if Y^* is not almost surely a minimum point of $h(\cdot)$, namely, $P\{c < h(Y^*)\} > 0$, then there exists $z = (z_1, \cdots, z_n) \in Z$ such that $P\{\bar{h}(z) < h(Y^*)\} > 0$. It is easy to see then that there is $y^* \in \mathbf{Q}$ with $P\{h(y^*) < h(Y^*)\} > 0$. Put $A := \{\omega : h(y^*, \omega) < h(Y^*(\omega), \omega)\}$, and $Y' := y^* \mathbf{1}_A + Y^* \mathbf{1}_{A^c}$. Then $Y' \in L^q(\mathcal{F}, \mathbf{R})$, and $Eh(Y') < Eh(Y^*)$, leading to a contradiction. $\square$

Define a set-valued function G: $\cup_{x \in \mathbf{R}} \partial f(x) \to 2^{\mathbf{R}}$

$$G(y) := \{x \in \mathbf{R} : y \in \partial f(x)\}, \quad \forall y \in \cup_{x \in \mathbf{R}} \partial f(x),$$

and define g: $\cup_{x \in \mathbf{R}} \partial f(x) \to \mathbf{R}$ as the "inverse function" of ∂f as follows

$$g(y) := \text{argmin}_{x \in G(y)} |x|, \quad \forall y \in \cup_{x \in \mathbf{R}} \partial f(x).$$

In Appendix we prove that g is a well-defined function (on its domain), and the set of y's where $G(y)$ is not a singleton is countable. In other words, denoting

$$\Gamma := \{y \in \cup_{x \in \mathbf{R}} \partial f(x) : G(y) \text{ is a singleton}\},$$

then the set $[\cup_{x \in \mathbf{R}} \partial f(x)] \setminus \Gamma$ is countable. Moreover, $g(\cdot)$ is increasing on $\cup_{x \in \mathbf{R}} \partial f(x)$ and continuous at points in Γ (Proposition C.5).

The objective of this subsection is to identify the ranges of y_0 where problem (5.32) admits optimal solution(s) and, when it does, to obtain an optimal solution in various situations of f. It follows from Lemma 5.6.1 that problem (5.32) admits an optimal solution if and only if there exists a pair (λ, μ) satisfying the following condition:

$$\text{There is } Y^* \in L^q(\mathcal{F}, \mathbf{R}) \text{ with } Y^* \in G(\lambda - \mu\xi), \text{ a.s., } EY^* = 0, \text{ and } E[\xi Y^*] = y_0. \quad (5.34)$$

Moreover, when there exists a pair (λ, μ) satisfying the above condition, Y^* is one of the optimal solutions for (5.32). Remark that if $\mu \neq 0$, then, since the set $[\cup_{x \in \mathbf{R}} \partial f(x)] \setminus \Gamma$ is countable and the distribution of ξ has no atom, we have $P\{\lambda - \mu\xi \in \Gamma\} = 1$. In this case $G(\lambda - \mu\xi)$ is almost surely a singleton; hence problem (5.32) has a unique optimal solution $Y^* = g(\lambda - \mu\xi)$.

We will solve problem (5.32) in each of the following four (mutually exclusive) cases:

Case 1: The set $\cup_{x\in\mathbf{R}}\partial f(x)$ is upper bounded but not lower bounded;

Case 2: The set $\cup_{x\in\mathbf{R}}\partial f(x)$ is lower bounded but not upper bounded;

Case 3: $\cup_{x\in\mathbf{R}}\partial f(x) = \mathbf{R}$;

Case 4: The set $\cup_{x\in\mathbf{R}}\partial f(x)$ is both upper and lower bounded.

Let us first focus on Case 1. In this case, it follows from Proposition C.1 that $\cup_{x\in\mathbf{R}}\partial f(x)$ is either a closed interval $(-\infty, \bar{k}]$ or an open one $(-\infty, \bar{k})$ where

$$\bar{k} := \lim_{x\to+\infty} f'_+(x) \in \mathbf{R}. \tag{5.35}$$

It is also clear that $\lim_{y\to-\infty} g(y) = -\infty$. Moreover, in this case one only needs to consider $\mu \geq 0$ in searching for (λ, μ) satisfying condition (5.34), for otherwise $\cup_{x\in\mathbf{R}}\partial f(x)$ would be unbounded from above.

The following technical lemma plays an important role in the subsequent analysis.

Lemma 5.6.2 *In Case 1, assume that there are $\lambda_0 > f'_-(0), \mu_0 > 0$ such that $g(\lambda_0 - \mu_0\xi) \in L^q(\mathcal{F}, \mathbf{R})$. Then for any $\mu_1 \in (0, \mu_0), \lambda_1 \in (f'_-(0), \lambda_0)$, there exists $\gamma \in L^q(\mathcal{F}, \mathbf{R}^+)$, such that $|g(\lambda - \mu\xi)| \leq \gamma$ for any $\mu \in [0, \mu_1]$ and $\lambda \in [f'_-(0), \lambda_1]$. If in addition $\xi g(\lambda_0 - \mu_0\xi) \in L^q(\mathcal{F}, \mathbf{R})$, then γ satisfies $\xi\gamma \in L^q(\mathcal{F}, \mathbf{R})$.*

Proof: Since $g(\cdot)$ is increasing (Proposition C.5), for any $\mu \in [0, \mu_1], \lambda \in [f'_-(0), \lambda_1]$, we have

$$g(f'_-(0) - \mu_1\xi) \leq g(\lambda - \mu\xi) \leq g(\lambda_1).$$

On the other hand, on the set $\{\omega : \xi(\omega) \leq \frac{\lambda_0 - f'_-(0)}{\mu_0 - \mu_1}\}$, we have

$$g(f'_-(0) - \mu_1\xi) \geq g\left(\frac{\mu_0 f'_-(0) - \lambda_0\mu_1}{\mu_0 - \mu_1}\right);$$

and on the set $\{\omega : \xi(\omega) > \frac{\lambda_0 - f'_-(0)}{\mu_0 - \mu_1}\}$ we have

$$g(f'_-(0) - \mu_1\xi) \geq g(\lambda_0 - \mu_0\xi).$$

Thus, if we put

$$\gamma := g(\lambda_1) + \left|g\left(\frac{\mu_0 f'_-(0) - \lambda_0\mu_1}{\mu_0 - \mu_1}\right)\right| + |g(\lambda_0 - \mu_0\xi)|,$$

then γ meets the requirement. $\square$

Lemma 5.6.3 *In Case 1, for any given* $\lambda \in (-\infty, \bar{k})$*,* $g_\lambda(\mu) := Eg(\lambda - \mu\xi)$ *is strictly decreasing in* $\mu \in \mathbf{R}^+$*.*

Proof: Since $g(\cdot)$ is increasing, $g_\lambda(\cdot)$ is decreasing. Moreover, for any $\mu > 0$, $g_\lambda(\mu) < g_\lambda(0)$. Indeed, if $g_\lambda(\mu) = g_\lambda(0)$, then $Eg(\lambda - \mu\xi) = Eg(\lambda)$ leading to $g(\lambda - \mu\xi) = g(\lambda)$. This, in turn, implies that $\lambda - \mu\xi \in \partial f(g(\lambda))$ which contradicts to the boundedness of $\partial f(g(\lambda))$.

Next, for any $0 < \mu_1 < \mu_2$, if $g_\lambda(\mu_1) = g_\lambda(\mu_2)$, then $g(\lambda - \mu_1\xi) = g(\lambda - \mu_2\xi)$ a.s.. We are to show that in this case $g(\cdot)$ must be constant on $(-\infty, \lambda - 1]$. In fact, if $g(\cdot)$ is not constant on $(-\infty, \lambda - 1]$, then for any $\epsilon > 0$, there exists $y_1 \leq \lambda - 1$ such that $g(y_1) < g(y_1 + \epsilon)$. Take $\epsilon = (\mu_2 - \mu_1)/(2\mu_2)$. Then it is straightforward to verify that $\frac{\lambda - y_1}{\mu_2} < \frac{\lambda - (y_1 + \epsilon)}{\mu_1}$. Now, if $\xi \in [\frac{\lambda - y_1}{\mu_2}, \frac{\lambda - (y_1+\epsilon)}{\mu_1}]$, then the monotonicity of $g(\cdot)$ yields $g(\lambda - \mu_1\xi) \geq g(y_1 + \epsilon)$ and $g(\lambda - \mu_2\xi) \leq g(y_1)$. It then follows from the inequality $g(y_1) < g(y_1 + \epsilon)$ that $P\{g(\lambda - \mu_2\xi) < g(\lambda - \mu_1\xi)\} \geq P\{\xi \in [\frac{\lambda - y_1}{\mu_2}, \frac{\lambda - (y_1+\epsilon)}{\mu_1}]\} > 0$, which contradicts the assumption that $g(\lambda - \mu_1\xi) = g(\lambda - \mu_2\xi)$ a.s..

We have shown that $g(\cdot)$ is constant on $(-\infty, \lambda - 1]$; nevertheless this is impossible because $\lim_{y\to-\infty} g(y) = -\infty$. The proof is complete. $\square$

Theorem 5.6.1 *In Case 1, assume that there are* $\lambda_0 > f'_-(0), \mu_0 > 0$ *such that* $g(\lambda_0 - \mu_0\xi) \in L^q(\mathcal{F}, \mathbf{R})$ *and* $Eg(\lambda_0 - \mu_0\xi) = 0$*. Then for any* $\lambda \in [f'_-(0), \lambda_0]$*, there exists a unique* $0 \leq \mu(\lambda) \leq \mu_0$ *such that* $g(\lambda - \mu(\lambda)\xi) \in L^q(\mathcal{F}, \mathbf{R})$ *and* $Eg(\lambda - \mu(\lambda)\xi) = 0$*. Moreover,* $\mu(\lambda) = 0$ *for* $\lambda \in [f'_-(0), f'_+(0)] \equiv \partial f(0)$*, and* $\mu(\cdot)$ *is continuous and strictly increasing on* $[f'_+(0), \lambda_0]$*.*

Proof: For any fixed $\lambda \in (f'_-(0), \lambda_0)$, define $g_\lambda(\mu) := Eg(\lambda - \mu\xi)$ for $\mu \in [0, \mu_0)$. It follows from Lemma 5.6.2 that for any $\mu_1 \in (0, \mu_0)$, the family of random variables $\{g(\lambda - \mu\xi) : \mu \in [0, \mu_1]\}$ are uniformly integrable. Hence by the dominated convergence theorem $g_\lambda(\cdot)$ is continuous on $[0, \mu_0)$. On the other hand, $g(\lambda - \mu\xi)$ is decreasing when $\mu \uparrow \mu_0$, and when $\mu_0 > \mu > \mu_0/2$, $g(\lambda - \mu\xi) \leq g(\lambda - \xi\mu_0/2) \in L^q(\mathcal{F}, \mathbf{R})$. Hence, the monotonic convergence theorem yields

$$\lim_{\mu \uparrow \mu_0} Eg(\lambda - \mu\xi) = E \lim_{\mu \uparrow \mu_0} g(\lambda - \mu\xi) = Eg(\lambda - \mu_0\xi).$$

Note that the above equality may take the value of $-\infty$. If $Eg(\lambda - \mu_0\xi) > -\infty$, then the strict monotonicity of g leads to $Eg(\lambda - \mu_0\xi) < Eg(\lambda_0 - \mu_0\xi) = 0$. Thus it always holds that $\lim_{\mu\uparrow\mu_0} g_\lambda(\mu) < 0$. But $g_\lambda(0) \equiv Eg(\lambda) \geq Eg(f'_-(0)) = 0$; so it follows from the facts that $g_\lambda(\cdot)$ is *strictly* decreasing (Lemma 5.6.3) and continuous on $[0, \mu_0)$ that there exists a unique $\mu(\lambda) \in [0, \mu_0)$ with $g_\lambda(\mu(\lambda)) \equiv Eg(\lambda - \mu(\lambda)\xi) = 0$. Moreover, Lemma 5.6.2 ensures that $g(\lambda - \mu(\lambda)\xi) \in L^q(\mathcal{F}, \mathbf{R})$.

To prove the second part of the theorem, first notice that $\lambda_0 > f'_+(0)$. Indeed, if it is not true, then $\lambda_0 \in \partial f(0)$ and hence $g(\lambda_0) = 0$. However, appealing to Lemma 5.6.3 we have $Eg(\lambda_0 - \mu_0\xi) > g(\lambda_0) = 0$ which is a contradiction. Now, whenever $\lambda \in [f'_-(0), f'_+(0)] \equiv \partial f(0)$, we have $Eg(\lambda) = g(\lambda) = 0$; thus the uniqueness of $\mu(\lambda)$ yields $\mu(\lambda) = 0$. Next, consider $\lambda_0 \geq \lambda_1 > \lambda_2 \geq f'_+(0)$. Since $\mu(\lambda) > 0$ whenever $\lambda > f'_+(0)$, and $Eg(\lambda_2 - \mu\xi) < Eg(\lambda_1 - \mu\xi)$ whenever $\mu > 0$, we have $g_{\lambda_2}(\mu(\lambda_1)) \equiv Eg(\lambda_2 - \mu(\lambda_1)\xi) < Eg(\lambda_1 - \mu(\lambda_1)\xi) \equiv 0 \equiv g_{\lambda_2}(\mu(\lambda_2))$. Since $g_{\lambda_2}(\cdot)$ is strictly decreasing, we conclude $\mu(\lambda_1) > \mu(\lambda_2)$, proving that $\mu(\cdot)$ is strictly increasing on $[f'_+(0), \lambda_0]$.

Next we show by contradiction the right continuity of $\mu(\cdot)$ on $[f'_+(0), \lambda_0)$. Assume that there exists $\lambda \in [0, \lambda_0)$, and $\epsilon > 0$ such that for any $\lambda' > \lambda$, $\mu(\lambda') > \mu(\lambda) + \epsilon$. Without loss of generality, suppose $\mu(\lambda) + \epsilon < \mu(\lambda_0)$. Then

$$0 = \lim_{\lambda'\downarrow\lambda} Eg(\lambda' - \mu(\lambda')\xi) \leq \lim_{\lambda'\downarrow\lambda} Eg(\lambda' - (\mu(\lambda) + \epsilon)\xi).$$

On the other hand, it follows from Lemma 5.6.2 that the family of random variables $\{g(\lambda - (\mu(\lambda) + \epsilon)\xi) : \lambda \in [\lambda, \lambda_1]\}$, for any fixed $\lambda_1 \in (\lambda, \lambda_0)$, is uniformly integrable. Therefore we have

$$\lim_{\lambda'\downarrow\lambda} Eg(\lambda' - (\mu(\lambda) + \epsilon)\xi) = Eg(\lambda - (\mu(\lambda) + \epsilon)\xi) < Eg(\lambda - \mu(\lambda)\xi) = 0,$$

leading to a contradiction.

It finally remains to prove the left continuity of $\mu(\cdot)$ on $(f'_+(0), \lambda_0]$. Assume that there exists $\lambda \in (f'_+(0), \lambda_0]$ and $\epsilon > 0$ such that for any $\lambda' < \lambda$, $\mu(\lambda') < \mu(\lambda) - \epsilon$. Without loss of generality, suppose $\mu(\lambda) - \epsilon > 0$. Then

$$0 = \lim_{\lambda'\uparrow\lambda} Eg(\lambda' - \mu(\lambda')\xi) \geq \lim_{\lambda'\uparrow\lambda} Eg(\lambda' - (\mu(\lambda) - \epsilon)\xi).$$

Obviously, $g(\lambda' - (\mu(\lambda) - \epsilon)\xi)$ is increasing when $\lambda' \uparrow \lambda$, and when $\lambda' > \lambda/2$, $g(\lambda' - (\mu(\lambda) - \epsilon)\xi) \geq g(\lambda/2 - (\mu(\lambda) - \epsilon)\xi) \in L^q(\mathcal{F}, \mathbf{R})$ by virtue of Lemma 5.6.2. Hence by the monotonic convergence theorem,

$$\lim_{\lambda'\uparrow\lambda} Eg(\lambda' - (\mu(\lambda) - \epsilon)\xi) = Eg(\lambda - (\mu(\lambda) - \epsilon)\xi) > Eg(\lambda - \mu(\lambda)\xi) = 0.$$

Again, this is a contradiction. $\square$

Define

$$\begin{cases} \bar{\Lambda} := \{\lambda \in [f'_-(0), \bar{k}] : \text{There exists } \mu = \mu(\lambda) \text{ so that } g(\lambda - \mu(\lambda)\xi) \in L^q(\mathcal{F}, \mathbf{R}), \\ \qquad\qquad Eg(\lambda - \mu(\lambda)\xi) = 0, \;\; \xi g(\lambda - \mu(\lambda)\xi) \in L^1(\mathcal{F}, \mathbf{R})\}, \\ \bar{\lambda} := \sup_{\lambda \in \bar{\Lambda}} \lambda, \\ \tilde{g}(\lambda) := E[\xi g(\lambda - \mu(\lambda)\xi)], \;\; \lambda \in [f'_-(0), \bar{\lambda}). \end{cases} \tag{5.36}$$

Notice that $\bar{\Lambda} \neq \emptyset$, since $\partial f(0) \subseteq \bar{\Lambda}$. As a result $f'_+(0) \leq \bar{\lambda} \leq \bar{k}$. Also, by virtue of Lemma 5.6.2 and Theorem 5.6.1, $[f'_-(0), \bar{\lambda}) \subseteq \bar{\Lambda}$.

Theorem 5.6.2 *In Case 1, $\tilde{g}(\lambda) = 0$ for $\lambda \in [f'_-(0), f'_+(0)] \equiv \partial f(0)$, and $\tilde{g}(\cdot)$ is continuous and strictly decreasing on $[f'_+(0), \bar{\lambda})$. Moreover, if $\bar{\lambda} \in \bar{\Lambda}$ and $\bar{\lambda} < \bar{k}$, then $\tilde{g}(\cdot)$ is also left continuous at $\bar{\lambda}$.*

Proof: Theorem 5.6.1 provides that $\mu(\lambda) = 0$ for any $\lambda \in \partial f(0)$; hence $\tilde{g}(\lambda) = E[\xi g(\lambda)] = 0$. Furthermore, for $\bar{\lambda} > \lambda_1 > \lambda_2 \geq f'_+(0)$ (if $\bar{\lambda} \in F$, then λ_1 may take the value of $\bar{\lambda}$), it follows from Theorem 5.6.1 that $\mu(\lambda_1) > \mu(\lambda_2) \geq 0$. Denote $\xi_0 := \frac{\lambda_1 - \lambda_2}{\mu(\lambda_1) - \mu(\lambda_2)} > 0$. If $\xi \geq \xi_0$, then $\lambda_1 - \mu(\lambda_1)\xi \leq \lambda_2 - \mu(\lambda_2)\xi$ resulting in $g(\lambda_1 - \mu(\lambda_1)\xi) - g(\lambda_2 - \mu(\lambda_2)\xi) \leq 0$. Similarly, if $\xi < \xi_0$, then $g(\lambda_1 - \mu(\lambda_1)\xi) - g(\lambda_2 - \mu(\lambda_2)\xi) \geq 0$. As a consequence,

$$\begin{aligned} & \tilde{g}(\lambda_1) - \tilde{g}(\lambda_2) \\ = \;& E\{\xi[g(\lambda_1 - \mu(\lambda_1)\xi) - g(\lambda_2 - \mu(\lambda_2)\xi)]\} \\ = \;& E\{\xi[g(\lambda_1 - \mu(\lambda_1)\xi) - g(\lambda_2 - \mu(\lambda_2)\xi)]\mathbf{1}_{\xi \geq \xi_0}\} + E\{\xi[g(\lambda_1 - \mu(\lambda_1)\xi) - g(\lambda_2 - \mu(\lambda_2)\xi)]\mathbf{1}_{\xi < \xi_0}\} \\ \leq \;& \xi_0 E\{[g(\lambda_1 - \mu(\lambda_1)\xi) - g(\lambda_2 - \mu(\lambda_2)\xi)]\mathbf{1}_{\xi \geq \xi_0}\} + \xi_0 E\{[g(\lambda_1 - \mu(\lambda_1)\xi) - g(\lambda_2 - \mu(\lambda_2)\xi)]\mathbf{1}_{\xi < \xi_0}\} \\ = \;& \xi_0 E[g(\lambda_1 - \mu(\lambda_1)\xi) - g(\lambda_2 - \mu(\lambda_2)\xi)] \\ = \;& 0. \end{aligned}$$

Moreover, if $\tilde{g}(\lambda_1) - \tilde{g}(\lambda_2) = 0$, then $g(\lambda_1 - \mu(\lambda_1)\xi) = g(\lambda_2 - \mu(\lambda_2)\xi)$ a.s.. By a reasoning similar to that in the proof of Lemma 5.6.3, we can prove that this is impossible. So $\tilde{g}(\cdot)$ is strictly decreasing on $[f'_+(0), \bar{\lambda})$.

Fix $\lambda \in [f'_+(0), \bar{\lambda})$. There is $\lambda_0 \in \bar{\Lambda}$ with $\lambda < \lambda_0$. By Lemma 5.6.2, the family $\{\xi g(\lambda' - \mu(\lambda')\xi) : \lambda' \in [0, (\lambda + \lambda_0)/2]\}$ is uniformly integrable. Thus by the continuity of $\mu(\cdot)$, we have

$$\begin{aligned}\lim_{\lambda' \to \lambda} \tilde{g}(\lambda') &= \lim_{\lambda' \to \lambda} E[\xi g(\lambda' - \mu(\lambda')\xi)] \\ &= E[\lim_{\lambda' \to \lambda} \xi g(\lambda' - \mu(\lambda')\xi)] \\ &= E[\xi g(\lambda - \mu(\lambda)\xi)] \\ &= \tilde{g}(\lambda).\end{aligned}$$

This proves the continuity of $\tilde{g}(\cdot)$ on $[f'_+(0), \bar{\lambda})$.

Finally, in the case when $\bar{\lambda} \in \bar{\Lambda}$ and $\bar{\lambda} < \bar{k}$, one has

$$\begin{aligned}\tilde{g}(\bar{\lambda}) &\leq \lim_{\lambda' \uparrow \bar{\lambda}} \tilde{g}(\lambda') \\ &= \lim_{\xi\lambda' \uparrow \bar{\lambda}} E[g(\lambda' - \mu(\lambda')\xi)] \\ &\leq \lim_{\lambda' \uparrow \bar{\lambda}} E[\xi g(\bar{\lambda} - \mu(\lambda')\xi)].\end{aligned}$$

On the other hand, since $g(\cdot)$ is increasing, we have $|g(\bar{\lambda} - \mu(\lambda')\xi)| \leq |g(\bar{\lambda})| + |g(\bar{\lambda} - \mu(\bar{\lambda})\xi)|$. Thus the dominated convergence theorem yields

$$\begin{aligned}\lim_{\lambda' \uparrow \bar{\lambda}} E[g(\bar{\lambda} - \mu(\lambda')\xi)\xi] &= E[g(\bar{\lambda} - \mu(\bar{\lambda})\xi)] \\ &= \tilde{g}(\bar{\lambda}).\end{aligned}$$

Therefore, $\tilde{g}(\cdot)$ is left continuous at $\bar{\lambda}$. $\square$

The following result gives the complete solution to problem (5.32) for Case 1.

Theorem 5.6.3 *Consider Case 1.*

(i) *If $\bar{\lambda} \notin \bar{\Lambda}$, then (5.32) admits an optimal solution if and only if $y_0 \in (\underline{y}, 0]$, where $\underline{y} = \lim_{\lambda \uparrow \bar{\lambda}} \tilde{g}(\lambda)$. If $\bar{\lambda} \in \bar{\Lambda}$, then (5.32) admits an optimal solution if and only if $y_0 \in \{\tilde{g}(\bar{\lambda})\} \cup (\underline{y}, 0]$. If in addition $\bar{\lambda} < \bar{k}$, then $\tilde{g}(\bar{\lambda}) = \underline{y}$.*

(ii) *When $y_0 = 0$, $Y^* := 0$ is the unique optimal solution to (5.32).*

(iii) *When $y_0 < 0$ and the existence of optimal solution is assured, $Y^* := g(\lambda - \mu(\lambda)\xi)$ is the unique optimal solution to (5.32), and the optimal value of the objectives is $Ef^*(\lambda - \mu(\lambda)\xi)$, where f^* is the conjugate function of f defined by $f^*(y) = \inf_{x\in\mathbf{R}}\{f(x) - xy\}$ and λ is the unique solution to $\tilde{g}(\lambda) = y_0$.*

Proof: (i) The "if" part follows immediately from Theorem 5.6.2. To prove the "only if" part, suppose that (5.32) admits an optimal solution Y^*, then there exists a pair (λ, μ) satisfying condition (5.34). If $\lambda < f'_-(0)$, then $\mu = 0$ (for otherwise $Eg(\lambda - \mu\xi) < Eg(\lambda) \leq g(f'_-(0)) = 0$). Hence it follows from (5.34) that $EY^* = 0$ and $Y^* \in G(\lambda)$, a.s. or $\lambda \in \partial f(Y^*)$, a.s.. If $P(Y^* = 0) < 1$, then $P(Y^* > 0) > 0, P(Y^* < 0) > 0$. Therefore $\lambda \in [\cup_{x>0}\partial f(x)] \cap [\cup_{x<0}\partial f(x)]$, which is impossible by Proposition C.2 and the fact that f is strictly convex at 0. Thus $P(Y^* = 0) = 1$ and, consequently, $y_0 = E[\xi Y^*] = 0$. On the other hand, if $\lambda \geq f'_-(0)$, then the conclusion follows from Theorem 5.6.2.

(ii) If $y_0 = 0$, it follows from Jensen's inequality that, for any feasible solution Y of (5.32), $Ef(Y) \geq f(EY) = f(0) \equiv Ef(0)$. Hence $Y^* := 0$ is an optimal solution. To prove that Y^* is the only solution, let Y be any feasible solution of (5.32) with $P(Y \neq 0) > 0$. Since f is strictly convex at 0, there exists an affine function $g(x) = ax + b$ so that $f(0) = g(0)$ and $f(x) > g(x) \quad \forall x \neq 0$. Therefore $P(f(Y) > g(Y)) > 0$, resulting in $Ef(Y) > Eg(Y) = g(EY) = g(0) = f(0) = Ef(0)$. This shows that Y is not optimal.

(iii) From Theorem 5.6.2, we can easily see $Y^* = g(\lambda - \mu(\lambda)\xi)$ is the unique optimal solution to problem 5.32. and the optimal value of the objective is $Ef(Y^*) = Ef(g(\lambda - \mu(\lambda)\xi)) = Ef^*(\lambda - \mu(\lambda)\xi)$. $\square$

Note that the "if" part of Theorem 5.6.3-(i) does not require the strict convexity of f at 0. However, this assumption can*not* be dropped for the "only if" part; see the following example.

Example 5.6.1 Take $f(x) = (x^2 - 1)\mathbf{1}_{x<-1}$, which is *not* strictly convex at 0. It is easy to see that $\cup_{x\in\mathbf{R}}\partial f(x) = (-\infty, 0]$. Pick $\alpha \in \mathbf{R}$ such that $P(\xi > \alpha) > \frac{1}{2} > P(\xi \leq \alpha) > 0$, and take $Y^* := \frac{P(\xi\leq\alpha)}{P(\xi>\alpha)}\mathbf{1}_{\xi>\alpha} - \mathbf{1}_{\xi\leq\alpha}$. Then, $EY^* = 0$, and $y_0^* := E[\xi Y^*] = P(\xi \leq \alpha)[E(\xi|\xi > \alpha) - E(\xi|\xi \leq \alpha)] > 0$. On the other hand, $Y^* \geq -1$ a.s., hence $Ef(Y^*) = 0$. This shows that problem (5.32) does admit an optimal solution Y^* even though $y_0 = y_0^* > 0$.

We have now completed the study on Case 1. As for Case 2, it can be turned into Case 1 by considering $\tilde{f}(x) = f(-x)$. Hence we only state the result.

Set

$$\underline{k} := \lim_{x \to -\infty} f'_-(x) \in \mathbf{R}, \tag{5.37}$$

and define

$$\begin{cases} \underline{\Lambda} := \{\lambda \in [\underline{k}, f'_+(0)] : \text{There exists } \mu = \mu(\lambda) \text{ so that } g(\lambda - \mu(\lambda)\xi) \in L^q(\mathcal{F}, \mathbf{R}), \\ \qquad\qquad Eg(\lambda - \mu(\lambda)\xi) = 0,\ \ \xi g(\lambda - \mu(\lambda)\xi) \in L^1(\mathcal{F}, \mathbf{R})\}, \\ \underline{\lambda} := \inf_{\lambda \in \underline{\Lambda}} \lambda, \\ \tilde{g}(\lambda) := E[\xi g(\lambda - \mu(\lambda)\xi)],\ \ \lambda \in (\underline{\lambda}, f'_+(0)]. \end{cases} \tag{5.38}$$

Theorem 5.6.4 *Consider Case 2.*

(i) *If $\underline{\lambda} \notin \underline{\Lambda}$, then (5.32) admits an optimal solution if and only if $y_0 \in [0, \bar{y})$, where $\bar{y} = \lim_{\lambda \downarrow \underline{\lambda}} \tilde{g}(\lambda)$. If $\underline{\lambda} \in \underline{\Lambda}$, then (5.32) admits an optimal solution if and only if $y_0 \in \{\tilde{g}(\lambda)\} \cup [0, \bar{y})$. If in addition $\underline{\lambda} > \underline{k}$, then $\tilde{g}(\bar{\lambda}) = \bar{y}$.*

(ii) *When $y_0 = 0$, $Y^* := 0$ is the unique optimal solution to (5.32).*

(iii) *When $y_0 > 0$ and the existence of optimal solution is assured, and the optimal value of the objectives is $Ef^*(\lambda - \mu(\lambda)\xi)$, where f^* is the conjugate function of f defined by $f^*(y) = \inf_{x \in \mathbf{R}}\{f(x) - xy\}$ and λ is the unique solution to $\tilde{g}(\lambda) = y_0$.*

Let us now turn to Case 3. It can be dealt with similarly combining the analysis for the previous two cases. Define

$$\begin{cases} \Lambda := \{\lambda \in \mathbf{R} : \text{There exists } \mu = \mu(\lambda) \text{ so that } g(\lambda - \mu(\lambda)\xi) \in L^q(\mathcal{F}, \mathbf{R}), \\ \qquad\qquad Eg(\lambda - \mu(\lambda)\xi) = 0,\ \ \xi g(\lambda - \mu(\lambda)\xi) \in L^1(\mathcal{F}, \mathbf{R})\}, \\ \bar{\lambda} := \sup_{\lambda \in \Lambda} \lambda,\ \ \underline{\lambda} := \inf_{\lambda \in \Lambda} \lambda, \\ \tilde{g}(\lambda) := E[\xi g(\lambda - \mu(\lambda)\xi)],\ \ \lambda \in (\underline{\lambda}, \bar{\lambda}), \\ \bar{y} := \lim_{\lambda \downarrow \underline{\lambda}} \tilde{g}(\lambda),\ \ \ \underline{y} := \lim_{\lambda \uparrow \bar{\lambda}} \tilde{g}(\lambda). \end{cases} \tag{5.39}$$

Theorem 5.6.5 *Consider Case 3. Problem (5.32) admits an optimal solution if and only if $y_0 \in A \cup B$, where*

$$A = \begin{cases} [\underline{y}, 0], & \text{if } \bar{\lambda} \in \Lambda \\ (\underline{y}, 0], & \text{if } \bar{\lambda} \notin \Lambda \end{cases}, \qquad B = \begin{cases} [0, \bar{y}], & \text{if } \underline{\lambda} \in \Lambda \\ [0, \bar{y}), & \text{if } \underline{\lambda} \notin \Lambda. \end{cases}$$

Moreover, when $y_0 = 0$*,* $Y^* := 0$ *is the unique optimal solution to (5.32), and when* $y_0 \neq 0$ *and the existence of optimal solution is assured,* $Y^* := g(\lambda - \mu(\lambda)\xi)$ *is the unique optimal solution to (5.32), and the optimal value of the objectives is* $Ef^*(\lambda - \mu(\lambda)\xi)$*, where* f^* *is the conjugate function of* f *defined by* $f^*(y) = \inf_{x \in \mathbf{R}}\{f(x) - xy\}$ *and* λ *is the unique solution to* $\tilde{g}(\lambda) = y_0$*.*

The final case, Case 4, only has a trivial solution, as shown in the following theorem.

Theorem 5.6.6 *Consider Case 4. Problem (5.32) admits an optimal solution if and only if* $y_0 = 0$*, in which case the unique optimal solution is* $Y^* = 0$*.*

Proof: Suppose that Y^* is optimal to (5.32). Then there exists (λ, μ) so that $\lambda - \mu\xi \in \partial f(Y^*)$, a.s.. It follows from the uniform boundedness of $\partial f(x)$ that $\mu = 0$. Employing the same argument as in the proof of Theorem 5.6.3-(i) we conclude that $Y^* = 0$, a.s.. $\square$

5.6.2 Asymptotic optimal portfolios

As shown in Theorems 5.6.3, 5.6.4, and5.6.5, when y_0 is in the certain range, problem (5.32) admits a unique optimal solution; when y_0 is out of the range, (5.32) admits no optimal solution. For the latter case, our results provide no resolution for the problem.

As we saw at the beginning of this section, the value of the risk admits a lower bounded $Ef(Y) \geq f(EY) = f(0)$. So for any y_0, there exists a finite infimum for the risk. Obviously, $f(0)$ is not the infimum in general. What is the infimum? How to approach it? We will discuss these problems in this subsection.

In addition to the assumptions imposed in the last subsection on the random variable ξ, we suppose

Assumption 5.6.2 $\xi \in L^{q/(q-1)}(\mathcal{F}, \mathbf{R})$.

With this assumption, we claim

Theorem 5.6.7 *When* $f(y) = |y|^q$*, Problem (5.32) admits a unique optimal solution for any* $y_0 \in \mathbf{R}$*.*

Proof: Obviously $f(\cdot)$ is strictly convex, $f'(y) = q\mathrm{sgn}(y)|y|^{q-1}$. By the definition, $g(x) = \mathrm{sgn}(x)(|x|/q)^{1/(q-1)}$.

For any $\lambda > 0$, it is easy to verify that there exists $\mu(\lambda)$ such that $Eg(\lambda - \mu(\lambda)\xi) = 0$. Furthermore, for any λ, μ, $g(\lambda - \mu\xi) \in L^q(\mathcal{F}, \mathbf{R})$, $|\xi g(\lambda - \mu\xi)| = \xi(|\lambda - \mu\xi|/q)^{1/(q-1)} \leq q^{-1/(q-1)}|\lambda|^{1/(q-1)}\xi + q^{-1/(q-1)}|\mu|^{1/(q-1)}\xi^{q/(q-1)} \in L^1(\mathcal{F}, \mathbf{R})$. So $\bar{\lambda} = +\infty, \underline{\lambda} = -\infty$. By Theorem 5.6.5, the problem admits a unique solution for any $y_0 \in \mathbf{R}$. □

As we have shown in the last subsection, in problem (5.32), y_0 cannot be any number in order to assure the existence of optimal solutions. In the following, we will design a sequence of problems, which ensure the existence of optimal solutions for any y_0, and meanwhile converge to problem (5.32) in some sense.

For any $\alpha > 0$, define $f_\alpha(x) = f(x) + \alpha x^q$, and consider the following problem:

$$\begin{array}{ll} \text{Minimize} & Ef_\alpha(Y), \\ \text{subject to} & \begin{cases} EY = 0, \\ E[\xi Y] = y_0, \\ Y \in L^q(\mathcal{F}, \mathbf{R}), \end{cases} \end{array} \tag{5.40}$$

where $f(\cdot)$ is the one in problem (5.32).

Theorem 5.6.8 *For any convex function $f(\cdot)$, any $\alpha > 0$, Problem (5.40) admits a unique optimal solution for any $y_0 \in \mathbf{R}$.*

Proof: Obviously, $f_\alpha(y) = f(y) + \alpha y^q$ is strictly convex. Denote $g(x)$ as the inverse function of $\partial f_\alpha(\cdot)$. By the property that $\partial f(x)$ is bounded for any $x \in \mathbf{R}$, $g(+\infty) = +\infty, g(-\infty) = -\infty$. Furthermore, thanks to the fact that $\partial f_\alpha(x) = \alpha q \text{sgn}(x)|x|^{q-1} + \partial f(x)$ and $0 \in \partial f(0)$, we know $0 \leq g(x) \leq (x/(\alpha q))^{1/(q-1)}$, $\forall x > 0$, and $0 \geq g(x) \geq -(-x/(\alpha q))^{1/(q-1)}$, $\forall x < 0$.

For any λ, it is easy to prove that there exists $u(\lambda) \in \mathbf{R}$ such that $Eg(\lambda - \mu(\lambda)\xi) = 0$. For any $\lambda, \mu \in \mathbf{R}$, thanks to the property of $g(x)$ and Assumption 5.6.2, we can prove that $g(\lambda - \mu(\lambda)\xi) \in L^q(\mathcal{F}, \mathbf{R}), \xi g(\lambda - \mu(\lambda)\xi) \in L^1(\mathcal{F}, \mathbf{R})$. Therefore $\bar{\lambda} = +\infty, \underline{\lambda} = -\infty$.

Like in the previous subsection, denote $\tilde{g}(\lambda) = E[\xi g(\lambda - \mu(\lambda)\xi)]$. Now we need to prove that $\lim_{\lambda \to +\infty} \tilde{g}(\lambda) = -\infty, \lim_{\lambda \to -\infty} \tilde{g}(\lambda) = +\infty$. Here we only prove the former, as a proof of the latter is the same. It is easy to verify that $\lim_{\lambda \to +\infty} \mu(\lambda) = +\infty, \lim_{\lambda \to +\infty} \mu(\lambda) = +\infty$. As shown in the last subsection, $\mu(\lambda)$ is increasing in λ, and $\tilde{g}(\lambda)$ is decreasing in λ.

Now we consider the problem in two cases, $\underline{\lim}_{\lambda\to+\infty}\frac{\lambda}{\mu(\lambda)}=+\infty$ and $\underline{\lim}_{\lambda\to+\infty}\frac{\lambda}{\mu(\lambda)}<+\infty$, respectively. For the first case, take a sequence $\lambda_n\uparrow+\infty$ such that $\lim_{n\to+\infty}\frac{\lambda_n}{\mu(\lambda_n)}=+\infty$. For written convenience, denote $\mu_n=\mu(\lambda_n)$. Then

$$\begin{aligned}
\tilde{g}(\lambda_n) &= E[\xi g(\lambda_n-\mu_n\xi)] \\
&= E[(\xi-\lambda_n/\mu_n)g(\lambda_n-\mu_n\xi)] \\
&= E[(\xi-\lambda_n/\mu_n)g(\lambda_n-\mu_n\xi)\mathbf{1}_{\xi<\lambda_n/\mu_n}]+E[(\xi-\lambda_n/\mu_n)g(\lambda_n-\mu_n\xi)\mathbf{1}_{\xi\geq\lambda_n/\mu_n}] \\
&\leq E[(\xi-\lambda_n/\mu_n)g(\lambda_n-\mu_n\xi)\mathbf{1}_{\xi<\lambda_n/\mu_n}] \\
&\leq E[(\xi-\lambda_n/\mu_n)g(\lambda_n-\mu_n\xi)\mathbf{1}_{\xi<\lambda_n/(2\mu_n)}] \\
&\leq E[(\xi-\lambda_n/\mu_n)g(\lambda_n/2)\mathbf{1}_{\xi<\lambda_n/(2\mu_n)}] \\
&\leq E[(\xi-\lambda_n/(2\mu_n))g(\lambda_n/2)\mathbf{1}_{\xi<\lambda_n/(2\mu_n)}] \\
&= -g(\lambda_n/2)E[(\lambda_n/(2\mu_n)-\xi)^+] \\
&\to -\infty.
\end{aligned}$$

By the decreasing property, we know $\lim_{\lambda\to+\infty}\tilde{g}(\lambda)=-\infty$. For the second case. take a sequence $\lambda_n\uparrow+\infty$ such that $\frac{\lambda_n}{\mu(\lambda_n)}$ converges to a finite constant c. Denote $\mu_n=\mu(\lambda_n)$. Then

$$\begin{aligned}
\tilde{g}(\lambda_n) &= E[\xi g(\lambda_n-\mu_n\xi)] \\
&= E[(\xi-\lambda_n/\mu_n)g(\lambda_n-\mu_n\xi)\mathbf{1}_{\xi<\lambda_n/\mu_n}]+E[(\xi-\lambda_n/\mu_n)g(\lambda_n-\mu_n\xi)\mathbf{1}_{\xi\geq\lambda_n/\mu_n}] \\
&\leq E[(\xi-\lambda_n/\mu_n)g(\lambda_n-\mu_n\xi)\mathbf{1}_{\xi\geq\lambda_n/\mu_n}] \\
&\leq E[(\xi-\lambda_n/\mu_n)g(\lambda_n-\mu_n\xi)\mathbf{1}_{\xi\geq 2\lambda_n/\mu_n}] \\
&\leq E[(\xi-\lambda_n/\mu_n)g(-\lambda_n)\mathbf{1}_{\xi\geq 2\lambda_n/\mu_n}] \\
&\leq g(-\lambda_n)E[(\xi-\lambda_n/\mu_n)\mathbf{1}_{\xi\geq 2\lambda_n/\mu_n}] \\
&\to -\infty.
\end{aligned}$$

Therefore $\lim_{\lambda\to+\infty}\tilde{g}(\lambda)=-\infty$.

It follows then from Theorem 5.6.5 that problem (5.40) admits a unique solution for any $y_0\in\mathbf{R}$. □

Theorem 5.6.8 is very useful for studying problem (5.32) when it does not admit any optimal solution, in which case we can assure the existence of the optimal solution by adding a small term $\alpha|x|^q$ into the risk function. To assure that this approach works, we still need some convergence property of problem (5.40) when α converges to 0.

For any $\alpha > 0$, according to Theorem 5.6.8, problem (5.40) admits a unique optimal solution, denoted as Y_α.

Theorem 5.6.9 *$Ef(Y_\alpha)$ is increasing on* $(0, +\infty)$*.*

Proof: Given $0 < \alpha_1 < \alpha_2$, we have

$$\begin{aligned} E[f(Y_1) + \alpha_1|Y_1|^q] &\leq E[f(Y_2) + \alpha_1|Y_2|^q], \\ E[f(Y_2) + \alpha_2|Y_2|^q] &\leq E[f(Y_1) + \alpha_2|Y_1|^q]. \end{aligned}$$

Therefore $\alpha_2 E(|Y_2|^q - |Y_1|^q) \leq Ef(Y_1) - Ef(Y_2) \leq \alpha_1 E(|Y_2|^q - |Y_1|^q)$, which implies $E|Y_2|^q < E|Y_1|^q$, and so $Ef(Y_1) \leq Ef(Y_2)$. □

Theorem 5.6.10 *Denote by $H(y_0)$ the infimum of the risk function in problem (5.32). Then* $\lim_{\alpha\downarrow 0} Ef(Y_\alpha) = H(y_0)$*.*

Proof: It is easy to see that $Ef(Y_\alpha) \geq H(y_0)$. By the monotonicity of $Ef(Y_\alpha)$, we have $\lim_{\alpha\downarrow 0} Ef(Y_\alpha) \geq H(y_0)$.

For any $\epsilon > 0$, there exists a feasible Y such that $Ef(Y) < H(y_0) + \epsilon$. Take any $\alpha \in (0, \epsilon/E|Y|^q)$, then $E[f(Y_\alpha) + \alpha|Y_\alpha|^q] \leq E[f(Y) + \alpha|Y|^q] < H(y_0) + 2\epsilon$. So $Ef(Y_\alpha) < H(y_0) + 2\epsilon$. This implies $\lim_{\alpha\downarrow 0} Ef(Y_\alpha) = H(y_0)$ □

Remark 5.6.1 In fact, in the proof of Theorem 5.6.9, we can see that $E[f(Y_\alpha)+\alpha|Y_\alpha|^q]$ is increasing; and then in the proof of Theorem 5.6.10, we can see $\lim_{\alpha\downarrow 0} E[f(Y_\alpha)-\alpha|Y_\alpha|^q] = H(y_0)$. Remember that $E[f(Y_\alpha) - \alpha|Y_\alpha|^q] = Ef_\alpha^*(\lambda_\alpha - \mu_\alpha(\lambda_\alpha)\xi)$, where $f_\alpha^*(\cdot)$ is the conjugate function of $f_\alpha(\cdot)$, λ_α is the solution for $\tilde{g}_\alpha = y_0$, and $\mu_\alpha(\lambda)$ is the unique solution for $Eg_\alpha(\lambda - \mu\xi) = 0$.

Remark 5.6.2 When y_0 is not in the range specified in Theorems 5.6.3, 5.6.4, and 5.6.5, we can just take Y_α^* as the approximation optimal solution sequence to approach the infimum of the risk function.

5.6.3 The general mean–risk portfolio selection problem

Now we are in the position to solve the following continuous-time portfolio selection problem. For each parameter $z \in \mathbf{R}$:

$$\begin{aligned} &\text{Minimize} \quad Ef(x(T) - Ex(T)), \\ &\text{subject to} \quad \begin{cases} \pi(\cdot) \in L^2_{\mathcal{F}}(0,T;\mathbf{R}^m), \\ (x(\cdot),\pi(\cdot)) \text{ satisfies equation (2.5) with initial wealth } x_0, \\ Ex(T) = z, \end{cases} \end{aligned} \tag{5.41}$$

where $f : \mathbf{R} \to \mathbf{R}$ satisfies Assumption 5.6.1 and Assumption 5.6.2. In this model, the risk is measured by the expectation of a convex function value of the deviation $x(T) - Ex(T)$.

The associated static optimization problem of (5.41) is

$$\begin{aligned} &\text{Minimize} \quad Ef(X - z), \\ &\text{subject to} \quad \begin{cases} EX = z, \\ E[\rho(T)X] = x_0, \\ X \in L^2(\mathcal{F},\mathbf{R}), \end{cases} \end{aligned} \tag{5.42}$$

where $\rho \equiv \rho(T)$ is given by (5.2), satisfying Assumption 5.5.2. This problem is a special case of (5.32) with $Y = X - z,\ y_0 = x_0 - zE\rho,\ q = 2,\ \xi = \rho$. Hence the results in the previous subsection readily apply. In the following theorem, we use the same notation, such as $g(\cdot)$, $\bar{\Lambda}$, $\bar{\lambda}$, etc., as in the previous subsection where ξ is replaced by $\rho(T)$.

Theorem 5.6.11 *Under Assumptions 5.6.1 and atom, one has the following conclusions regarding the solution to the mean–risk portfolio selection problem (5.41):*

(i) *Assume that either $\cup_{x\in\mathbf{R}}\partial f(x) = (-\infty, \bar{k}]$ or $\cup_{x\in\mathbf{R}}\partial f(x) = (-\infty, \bar{k})$ for some $\bar{k} \in \mathbf{R}$. If $\bar{\lambda} \notin \bar{\Lambda}$, then (5.41) admits an optimal solution if and only if $x_0 - zE\rho \in (\underline{y}, 0]$, where $\underline{y} = \lim_{\lambda\uparrow\bar{\lambda}} \tilde{g}(\lambda)$. If $\bar{\lambda} \in \bar{\Lambda}$, then (5.41) admits an optimal solution if and only if $x_0 - zE\rho \in \{\tilde{g}(\bar{\lambda})\} \cup (\underline{y}, 0]$. If in addition $\bar{\lambda} < \bar{k}$, then $\tilde{g}(\bar{\lambda}) = \underline{y}$.*

(ii) *Assume that either $\cup_{x\in\mathbf{R}}\partial f(x) = [\underline{k}, \infty)$ or $\cup_{x\in\mathbf{R}}\partial f(x) = (\underline{k}, \infty)$ for some $\underline{k} \in \mathbf{R}$. If $\underline{\lambda} \notin \underline{\Lambda}$, then (5.41) admits an optimal solution if and only if $x_0 - zE\rho \in [0, \bar{y})$, where $\bar{y} = \lim_{\lambda\downarrow\underline{\lambda}} \tilde{g}(\lambda)$. If $\underline{\lambda} \in \underline{\Lambda}$, then (5.41) admits an optimal solution if and only if $x_0 - zE\rho \in \{\tilde{g}(\lambda)\} \cup [0, \bar{y})$. If in addition $\underline{\lambda} > \underline{k}$, then $\tilde{g}(\bar{\lambda}) = \bar{y}$.*

(iii) *Assume that* $\cup_{x\in\mathbf{R}}\partial f(x) = \mathbf{R}$. *Then (5.41) admits an optimal solution if and only if* $x_0 - zE\rho \in A \cup B$, *where*

$$A = \begin{cases} [\underline{y}, 0], & \text{if } \bar{\lambda} \in \Lambda \\ (\underline{y}, 0], & \text{if } \bar{\lambda} \notin \Lambda \end{cases}, \qquad B = \begin{cases} [0, \overline{y}], & \text{if } \underline{\lambda} \in \Lambda \\ [0, \overline{y}), & \text{if } \underline{\lambda} \notin \Lambda. \end{cases}$$

(iv) *Assume that there exists* $M_1, M_2 \in \mathbf{R}$ *such that* $\cup_{x\in\mathbf{R}}\partial f(x) \subset [M_1, M_2]$. *Then (5.41) admits an optimal solution if and only if* $z = x_0/E\rho$.

(v) *When* $z = x_0/E\rho$, *the optimal portfolio is the risk-free one. When* $z \neq x_0/E\rho$ *and the existence of optimal solution is assured, the unique optimal portfolio for (5.41) is given by*

$$\pi^*(t) = (\sigma(t)')^{-1}y^*(t), \tag{5.43}$$

where $(x^*(\cdot), y^*(\cdot))$ *is the unique solution to the BSDE*

$$\begin{cases} dx(t) = [r(t)x(t) + \theta(t)y(t)]dt + y(t)'dW(t) \\ x(T) = g(\lambda - \mu(\lambda)\rho) + z, \end{cases} \tag{5.44}$$

with λ *being the unique solution to* $E[\rho g(\lambda - \mu(\lambda)\rho)] = x_0 - zE\rho$.

(vi) *When problem (5.41) admits no optimal solution, the optimal value can be approached by the feasible solution sequence* $\{Y_\alpha : \alpha \downarrow 0\}$, *where* Y_α *is the unique optimal solution for problem (5.40) (with* $q = 2$*).*

5.6.4 Examples

In this subsection, we apply the general results obtained to several special problems.

Example 5.6.2 Let $f(x) = \alpha(x^+)^2 + \beta(x^-)^2$ with $\alpha, \beta > 0$. This corresponds to the weighted mean–variance model that has been studied in Section 5.3. f is strictly convex, $\cup_{x\in\mathbf{R}}\partial f(x) = \mathbf{R}$, and $g(y) = \frac{1}{2\alpha}y^+ - \frac{1}{2\beta}y^-$. For any $\lambda > 0$, it is straightforward to see that the equation $Eg(\lambda - \mu\rho) = 0$ has a unique solution $\mu(\lambda) = \lambda/\zeta$ where $\zeta > 0$ uniquely solves (5.12). Hence $\bar{\lambda} = +\infty$, and

$$\tilde{g}(\lambda) = \frac{E[\rho(\lambda - \mu(\lambda)\rho)^+]}{2\alpha} - \frac{E[\rho(\lambda - \mu(\lambda)\rho)^-]}{2\beta} = \lambda\tilde{g}(1).$$

As a result, $\lim_{\lambda\to+\infty}\tilde{g}(\lambda) = -\infty$ (recall that $\tilde{g}(1) < \tilde{g}(0) = 0$). Similarly, we can prove that $\underline{\lambda} = -\infty$ and $\lim_{\lambda\to-\infty}\tilde{g}(\lambda) = +\infty$. We can then apply Theorem 5.6.11-(iii) to conclude that the weighted mean-variance model admits a unique optimal solution for any $z \in \mathbf{R}$. Finally, the optimal portfolio obtained in Theorem 5.3.2 can be easily recovered by Theorem 5.6.11-(v). (It should be noted, however, Theorem 5.3.2 cannot be superseded as Assumption 5.5.2 is not imposed there.)

Example 5.6.3 Let $f(x) = (x^-)^2$. This is the mean–semivariance model investigated in Section 5.4. Clearly, f is convex, strictly convex at 0, and $\cup_{x\in\mathbf{R}}\partial f(x) = (-\infty, 0]$. The inverse function $g(y) = \frac{1}{2}y, \quad y \le 0$. It is easily seen that $\bar{\Lambda} = \{0\}$ and $\bar{\lambda} = 0 \in \bar{\Lambda}$. Now, $\tilde{g}(\lambda) = E[\rho g(\lambda - \mu\rho)] = \frac{1}{2}(E\rho - \frac{E\rho^2}{E\rho})\lambda$. Thus $\underline{y} = \lim_{\lambda\uparrow 0}\tilde{g}(\lambda) = 0$. It then follows from Theorem 5.6.11-(i) that the mean–semivariance model admits an optimal solution if and only if $z = x_0/E\rho$. (Again, this does not recover Theorem 5.4.1 completely due to Assumption 5.5.2.)

Example 5.6.4 Let $f(x) = |x|$. The corresponding portfolio selection problem is called the mean–absolute-deviation model. Single-period mean–absolute-deviation model is studied in [28]. Now, f is strictly convex at 0, and $\cup_{x\in\mathbf{R}}\partial f(x) = [-1, 1]$. Thus in view of Theorem 5.6.11-(iv) the continuous-time mean–absolute-deviation model admits an optimal solution if and only if $z = x_0/E\rho$, in which case the optimal portfolio is simply the risk-free one.

Example 5.6.5 Let $f(x) = e^{-x}$. This function captures the situation where lager deviation of the terminal wealth from its mean is heavily penalized. Again, f is strictly convex, $\cup_{x\in\mathbf{R}}\partial f(x) = (-\infty, 0)$ (hence $\bar{k} = 0$), and $g(y) = -\ln(-y),\ y < 0$. Now, the equation $Eg(0 - \mu\rho) = 0$ has a solution $\mu \equiv \mu(0) = e^{-E\ln\rho} > 0$, Moreover, $g(0 - \mu(0)\rho) = \int_0^T [r(s) + \frac{|\theta(s)|^2}{2}]ds + \int_0^T \theta(s)dW(s) + E\ln\rho \in L^2(\mathcal{F}, \mathbf{R})$. It follows then from Theorem 5.6.1 that $\bar{\Lambda} = [-1, 0]$ and, consequently, $\bar{\lambda} = 0 = \bar{k}$. Furthermore, $\tilde{g}(0) = E[g(0 - \mu(0)\rho)\rho] = (E\rho)(E\ln\rho) - E(\rho\ln\rho)$. On the other other hand, when $-1 < \lambda \uparrow 0$, $g(\lambda - \mu(\lambda)\rho) = -\ln(-\lambda + \mu(\lambda)\rho) \ge -\ln(1 + \mu(0)\rho) \ge -\mu(0)\rho$, and $g(\lambda - \mu(\lambda)\rho) = -\ln(-\lambda + \mu(\lambda)\rho) \le -\ln(\mu(\lambda)\rho) \le -\ln(\mu(-1/2)) - \ln\rho$. Thus the dominated convergence theorem ensures that $\underline{y} \equiv \lim_{\lambda\uparrow 0}\tilde{g}(\lambda) = \tilde{g}(0)$. By Theorem 5.6.11-(i), the mean–risk portfolio selection problem admits an optimal solution if and only if

$x_0 - zE\rho \in [(E\rho)(E\ln\rho) - E(\rho\ln\rho), 0]$ or, equivalently, $z \in [\frac{x_0}{E\rho}, \frac{x_0-(E\rho)(E\ln\rho)+E(\rho\ln\rho)}{E\rho}]$. Finally, by Theorem 5.6.11-(v), when the problem does admit an optimal solution, the optimal portfolio is the one replicating the claim $z - \ln(-\lambda + \mu\rho)$ where (λ, μ) is the unique solution pair to the following algebraic equation (which must admit a solution):

$$\begin{cases} E\ln(-\lambda + \mu\rho) = 0, \\ E[\rho\ln(-\lambda + \mu\rho)] = zE\rho - x_0. \end{cases}$$

Example 5.6.6 Let $f(x) = ((x-1)^-)^2$. The corresponding portfolio selection model is a variant of the mean–semivariance model, except that the terminal wealth being less than its mean plus 1 is now considered as risk. In this case, f is not strictly convex everywhere; but it is indeed strictly convex at 0. It is easy to see that $\cup_{x\in\mathbf{R}}\partial f(x) = (-\infty, 0]$ (hence $\bar{k} = 0$), and $g(y) = y/2 + 1,\ y \le 0$. Meanwhile the equation $Eg(0 - \mu\rho) = 0$ has a solution $\mu \equiv \mu(0) = 2/E\rho > 0$. By virtue of Theorem 5.6.1, $\bar{\Lambda} = [-2, 0]$ and, consequently, $\bar{\lambda} = 0 = \bar{k}$. Note that $g(0-\mu(0)\rho) = g(0-\mu\rho) = E\rho - E\rho^2/E\rho$, and $\underline{y} \equiv \lim_{\lambda\uparrow 0}\tilde{g}(\lambda) = \tilde{g}(0)$. By Theorem 5.6.11-(i) the original portfolio selection problem admits an optimal solution if and only if $x_0 - zE\rho \in [E\rho - E\rho^2/E\rho, 0]$ or, equivalently, $z \in [\frac{x_0}{E\rho}, \frac{x_0}{E\rho} + \frac{E\rho^2}{(E\rho)^2} - 1]$. At last, when the problem does admit an optimal solution, the optimal portfolio is the one replicating the claim $z + 1 + \frac{\lambda-\mu\rho}{2}$ where (λ, μ) is the unique solution pair to the following *linear* algebraic equation:

$$\begin{cases} \lambda - \mu E\rho = -2, \\ \lambda E\rho - \mu E\rho^2 = 2x_0 - 2(1+z)E\rho. \end{cases}$$

Compared with Example 5.6.3 it is interesting to see that a shift of the mean makes the mean–semivariance model, which does not admit an optimal solution in any non-trivial case, possess non-trivial optimal solution.

5.7 Fixed-target problem

In the mean-risk framework studied in Section 5.6, the risk is measured by the centered return based on its mean. Sometime, investors may look at some fixed target rather than the mean of the terminal return. In this section, we will investigate the fixed-target investment problem. For an instance, an investor want to track a certain index at time T with an initial endowment x_0.

We study the problem in the same market as in section 5.6, and adopt the $\rho(\cdot)$ defined in Section 5.2 and $\rho \stackrel{\triangle}{=} \rho(T)$. $Z \in L^2(\mathcal{F}_T, \mathbf{R})$ is a given target at time T for the investment, x_0 is a given initial wealth. We want to find a portfolio such that the terminal wealth is the closest to Z. Here we use $Ef(X-Y)$ to measure the "distance" from X to Y, where f is a convex real function, and $f(y) > f(0) = 0$ for any $y \in \mathbf{R}$. Then the fixed-target problem can be formulated as follows:

$$\begin{array}{ll} \text{Minimize} & Ef(x(T) - Z), \\ \text{subject to} & \begin{cases} \pi(\cdot) \in L^2_{\mathcal{F}}(0, T; \mathbf{R}^m), \\ (x(\cdot), \pi(\cdot)) \text{ satisfies equation (2.5) with initial wealth } x_0. \end{cases} \end{array} \tag{5.45}$$

Like before, thanks to the completeness of the market, we can fist solve out the optimal terminal wealth $x(T)$ of the problem (5.45) by its static problem:

$$\begin{array}{ll} \text{Minimize} & Ef(Y - Z), \\ \text{subject to} & \begin{cases} EY\rho = x_0. \\ Y \in L^2(\mathcal{F}_T, \mathbf{R}) \end{cases} \end{array} \tag{5.46}$$

We suppose

Assumption 5.7.1 *Assumption 5.5.2 hold; $f(\cdot)$ is a strictly convex function on $\mathbf{R}$, and $f(x) > f(0) = 0$, $\forall x \in \mathbf{R}$.*

By Proposition 3.4.1, we know (5.46) admits an optimal solution Y^* if and only if Y^* is feasible for problem (5.46) and there exists a $\lambda \in \mathbf{R}$ such that Y^* solves the following problem

$$\min_{Y \in L^2(\mathcal{F}_T, \mathbf{R})} E(f(Y - Z) - \lambda\rho Y) \tag{5.47}$$

By the same proof of Lemma 5.6.1, we conclude that $Y^* \in L^2(\mathcal{F}_T, \mathbf{R})$ is an optimal solution to problem (5.47) if and only

$$\lambda\rho \in \partial f(Y^* - Z), \quad P - \text{a.s.} \tag{5.48}$$

For any $y \in \cup_{x\in\mathbf{R}}\partial f(x)$, define $g(y) := \text{argmin}_{x\in\partial f(y)}|x|$. then g is continuous and strictly increasing on $\mathbf{R}\backslash\Gamma$, where $\Gamma = \{x \in \mathbf{R}, \partial f(x) \text{ is not a singleton } \}$ is a countable subset of $\mathbf{R}$.

Define $F_1 = \{\lambda \in \mathbf{R} : \lambda\rho \in \cup_{x\in\mathbf{R}}\partial f(x)$, then the following claim is straightforward.

Theorem 5.7.1

(i) If $\cup_{x\in\mathbf{R}}\partial f(x)$ *is lower bounded but not upper bounded, then* $F_1 = [0, +\infty)$*;*

(ii) If $\cup_{x\in\mathbf{R}}\partial f(x)$ *is upper bounded but not lower bounded, then* $F_1 = (-\infty, 0]$*;*

(iii) If $\cup_{x\in\mathbf{R}}\partial f(x)$ *is bounded on both sides, then* $F_1 = 0$*;*

(iv) If $\cup_{x\in\mathbf{R}}\partial f(x)$ *is unbounded on both sides, then* $F_1 = \mathbf{R}$*.*

Define $F_2 = \{\lambda \in F_1 : g(\lambda\rho) \in L^2(\mathcal{F}_T, \mathbf{R})\}$, then F_2 is a convex subset of $\mathbf{R}$. Furthermore, $\tilde{g}(\lambda) = E[\rho g(\lambda\rho)]$ is strictly increasing and continuous on F_2. Denote $F_3 = \{\tilde{g}(\lambda) + E[Z\rho] : \lambda \in F_2\}$, then F_3 is also a convex subset of F_2, therefore an interval on $\mathbf{R}$.

With these denotation, we have the following claim:

Theorem 5.7.2 *For any* $x_0 \in F_3$*, problem (5.46) admits a unique optimal solution* $Y^* = g(\lambda\rho) + Z$*, where* λ *is the unique solution for* $\tilde{g}(\lambda) = x_0 - E[Z\rho]$*; For any* $x_0 \notin F_3$*, problem (5.46) admits no optimal solution.*

Proof: The conclusion can be proved by the condition (5.47) and the fact that $\tilde{g}(\cdot)$ is continuous and strictly increasing function from F_2 to $F_3 - E[Z\rho]$. □

With Theorem 5.7.2 and the completeness of the financial market, the following theorem is straightforward.

Theorem 5.7.3 *For any* $x_0 \in F_3$*, problem (5.45) admits a unique optimal solution pair* $(x^*(\cdot), \pi^*(\cdot))$*, which is the solution pair for*

$$\begin{cases} dx(t) = r(t)x(t)dt + \pi(t)'[B(t)dt + \sigma(t)dW(t)] \\ x(T) = Y^* \end{cases} \tag{5.49}$$

where Y^* *is the one in Theorem 5.7.2. For any* $x_0 \notin F_3$*, problem (5.45) admits no optimal solution.*

Remark 5.7.1

(i) F_3 is nonempty, because $E[Z\rho] \in F_3$;

(ii) F_2 can be either closed or open at its extreme points. For example, Suppose $g^2(x) = \sum_{j=1}^{\infty} a_j x^j$, $\forall x \geq 0$, for some positive sequence $\{a_j\}$, and $g^2(x) < \infty$. Then $Eg^2(\lambda\rho) = E\sum_{j=1}^{\infty} a_j \lambda^j \rho^j$. By Tonelli Theorem,

$$Eg^2(\lambda\rho) = E\sum_{j=1}^{\infty} a_j \lambda^j \rho^j = \sum_{j=1}^{\infty} a_j E\rho^j \lambda^j = \sum_{j=1}^{\infty} b_j \lambda^j$$

where $b_j = a_j * E\rho^j$. If the power series $\sum_{j=1}^{\infty} b_j x^j$ converge at some $x_0 > 0$ but diverge at any $x > x_0$, then the right extreme point of F_2 is x_0 and $x_0 \in F_2$; If the power series $\sum_{j=1}^{\infty} b_j x^j$ converge at any $x < x_0$ but diverge at any $x \geq x_0$, then the right extreme point of F_2 is x_0 and $x_0 \notin F_2$.

As a concrete example, suppose ρ follows a exponential distribution with parameter λ_0, then $E\rho^j = \lambda_0^j j!$. Take $a_j = \frac{1}{j^2 E\rho^j} = \frac{1}{(j^2\lambda_0^j j!)}$, then $b_j = j^{-2}$. Therefore $Eg^2(1) = \sum_{j=1}^{\infty} J^{-2} < \infty$, but $Eg^2(x) = \infty$ for any $x > 1$. In this example, $g^2(x) = \sum_{j=1}^{\infty} a_j x^j$, it is finite for any $x \in \mathbf{R}^+$, and it's strictly increasing.

If we take $a_j = \frac{1}{j * E\rho^j}$, then we can easily find that the right extreme point of F_2 is 1 but $1 \notin F_2$.

We can construct the example for the left extreme point of F_2 in the same way.

(iii) When F_2 is open/closed at its right/left extreme point, F_3 is also open/closed at its right/left extreme point. This is because $\tilde{g}(\cdot)$ is continuous on F_2.

5.8 Mean-semivariance in single-period

In Section 5.4, we showed that in a continuous-time market, the mean-semivariance portfolio selection problem, save for a trivial case, admits no optimal solutions. This negative result was generalized to the mean-downside-risk portfolio selection problem in Section 5.5. In this section, we will show that, in sharp contrast, in a single-period market the mean-semivariance portfolio selection problem does admit optimal solutions in general.

Furthermore, the conclusion can be generalized to the mean-downside-risk framework with some additional conditions.

In this section, the notation is independent of that in the other sections.

Suppose there are $n\,(n \geq 2)$ securities traded in the market. The total return of the i^{th} security during the investment period is ξ_i (i.e., the payoff of \$1 for security i is \$$\xi_i$). Suppose $E\xi_i = r_i$ and $\mathrm{Var}(\xi_i) < +\infty$, and $\xi_1, \cdots, \xi_n$ are linearly independent, namely, $\alpha_1 = \alpha_2 = \cdots = \alpha_n = 0$ whenever $\sum_{i=1}^n \alpha_i \xi_i = 0$ for real number $\alpha_1, \cdots, \alpha_n$. The last assumption means that there are no redundant securities in the market.

The single-period mean-semivariance portfolio selection problem is as follows:

$$\begin{array}{ll} \text{Minimize} & E[(\sum_{i=1}^n x_i\xi_i - E\sum_{i=1}^n x_i\xi_i)^-]^2, \\ \text{subject to} & \begin{cases} \sum_{i=1}^n x_i = a, \\ \sum_{i=1}^n x_i r_i = z, \end{cases} \end{array} \tag{5.50}$$

where $x_i \in \mathbf{R}$ represents the capital amount invested in the i^{th} security, a is the amount of the initial investment, and $z \in \mathbf{R}$ is a given expectation level of the investment payoff.

Denote $R_i = \xi_i - r_i$. Then $ER_i = 0$ and we can rewrite the mean-semivariance problem (5.50) as

$$\begin{array}{ll} \text{Minimize} & E[(\sum_{i=1}^n x_i R_i)^-]^2, \\ \text{subject to} & \begin{cases} \sum_{i=1}^n x_i = a, \\ \sum_{i=1}^n x_i r_i = z. \end{cases} \end{array} \tag{5.51}$$

Replacing x_1 by the first constraint we get the following equivalent problem

$$\begin{array}{ll} \text{Minimize} & E[(aR_1 + \sum_{i=2}^n x_i(R_i - R_1))^-]^2, \\ \text{subject to} & \sum_{i=2}^n x_i(r_i - r_1) = z - ar_1. \end{array} \tag{5.52}$$

Before studying this problem, we introduce a useful lemma.

Lemma 5.8.1 *Let m be an integer, A, B_i be random variables with zero mean and finite variance, and $B = (B_1, \cdots, B_m)'$. If A, B are linearly independent, then the following optimization problem*

$$\min_{x \in I\!R^m} E[(A + B'x)^-]^2 \tag{5.53}$$

admits optimal solutions.

Proof: Define $S = \{(k,y) \in \mathbf{R}^{m+1} : 0 \le k \le 1, |y| = 1\}$, $l := \min_{(k,y)\in S} E[(kA + B'y)^-]^2$. When $(k,y) \in S$, $E[kA + B'y)^-]^2 \le E(kA + B'y)^2 \le 2(EA^2 + E|B|^2)$. Hence by the dominated convergence theorem, $E[(kA + B'y)^-]^2$ is continuous in (k,y) on the set S. Obviously, S is closed and bounded; so there exists $(k^*, y^*) \in S$ such that $l = E[(k^*A + B'y^*)^-]^2$.

We now prove $l > 0$ by contradiction. Obviously $l \ge 0$. If $l = 0$, then $k^*A + B'y^* = 0$. By the linear independence of A, B, we have $y^* = 0$, which contradicts to the fact that $|y^*| = 1$.

Next, for any $x \in \mathbf{R}^m$ with $|x| \ge 1$, we have

$$E[(A + B'x)^-]^2 = |x|^2 E[(\frac{A}{|x|} + B'\frac{x}{|x|})^-]^2 \ge |x|^2 l.$$

This shows that the function to be minimized in (5.53) is coercive (that is the function value tends to positive infinity when $|x|$ tends to positive infinity). As a result, the minimum, if any, must be within a certain closed bounded region. Hence (5.53) is effectively a problem of minimizing a continuous function over a closed bounded region, which therefore must admit optimal solutions. $\square$

Now let us return to problem (5.52), which is equivalent to the mean-semivariance portfolio selection problem (5.50). We study the problem in two cases. The first case is when all the securities have the same expected total return.

Theorem 5.8.1 *Suppose $r_i = r_1$ for all i. If $z \ne ar_1$, then problem (5.50) admits no feasible solution; if $z = ar_1$, then problem (5.50) admits optimal solutions.*

Proof: The first claim is obvious. So we assume $z = ar_1$. In this case, the constraint of problem (5.52) is satisfied automatically. Hence (5.52) becomes

$$\min_{x_i \in I\!R} E[(aR_1 + \sum_{i=2}^{n} x_i(R_i - R_1))^-]^2.$$

If $a = 0$, then $z = 0$, in which case $x_1 = \cdots = x_n = 0$ is obviously an optimal solution for problem (5.50). Suppose $a \ne 0$, then by the linear independence of $R_1, \cdots, R_n$, we know $aR_1, R_2 - R_1, \cdots, R_n - R_1$ are also linearly independent. By Lemma 5.8.1, problem (5.52) admits optimal solutions; so does the mean-semivariance portfolio selection problem (5.50). $\square$

The second case is when there exists i so that $r_i \neq r_1$. Without loss of generality, we suppose $r_2 \neq r_1$.

Theorem 5.8.2 *Suppose $r_2 \neq r_1$. Then for any $z \in \mathbf{R}$, problem (5.50) admits optimal solutions.*

Proof: In problem (5.52), replacing x_2 by $x_2 = \frac{z-ar_1}{r_2-r_1} - \sum_{i=3}^n x_i \frac{r_i-r_1}{r_2-r_1}$, we get the following equivalent problem:

$$\min_{(x_3,\cdots,x_n)\in I\!R^{n-2}} E[(aR_1+\frac{z-ar_1}{r_2-r_1}(R_2-R_1)+\sum_{i=3}^n x_i(R_i-R_1-(r_i-r_1)\frac{R_2-R_1}{r_2-r_1}))^-]^2. \quad (5.54)$$

Define $A = aR_1 + \frac{z-ar_1}{r_2-r_1}(R_2-R_1), B_i = R_i - R_1 - (r_i - r_1)\frac{R_2-R_1}{r_2-r_1}, B = (B_3, \cdots, B_n)'$, then $EA = 0$ and $EB = 0$. Suppose $A \neq 0$, then by the linear independence of $R_1, \cdots, R_n$, A, B are linearly independent. By Lemma 5.8.1, problem (5.52) admits optimal solutions, so does the mean-semivariance portfolio selection problem (5.50).

If $A = 0$, then $x_3 = \cdots = x_n = 0$ is obviously an optimal solution for problem (5.54), and therefore problem (5.50) admits at least an optimal solution. $\square$

In conclusion, we claim the existence of optimal portfolios for the mean-semivariance portfolio selection problem in the single-period market as follows.

Theorem 5.8.3 *Problem (5.50) admits optimal solutions if and only if it admits feasible solutions.*

Remark 5.8.1 The conclusion in Theorem 5.8.3 can be generalized to the following mean-downside-risk problem

$$\begin{array}{ll} \text{Minimize} & Ef(\sum_{i=1}^n x_i\xi_i - E\sum_{i=1}^n x_i\xi_i), \\ \text{subject to} & \begin{cases} \sum_{i=1}^n x_i = 1, \\ \sum_{i=1}^n x_i r_i = z, \end{cases} \end{array} \quad (5.55)$$

where $f(\cdot)$ is a real function satisfying:

(i) $f(x) = 0$ when $x \geq 0$;

(ii) $f(\cdot)$ is continuous, and decreasing on R^-;

(iii) there exists $q > 0$ and $L > 0$ such that $f(kx) \geq Lk^q f(x) \forall k > 0$;

(iv) $Ef(\sum_{i=1}^{n} x_i R_i)$ is continuous in $(x_1, \cdots, x_n)$.

The proof is the same as that of the mean-semivariance problem.

Remark 5.8.2 Compare the conclusions in this section with those in Section 5.4 and 5.5, we can find that the continuous-time market is essentially different from the discrete-time market.

5.9 Conclusion

In this chapter we have first solved a weighted mean–variance portfolio selection model in a complete continuous-time financial market. Inspired by its result, we have proved that, other than a trivial case, the mean–semivariance problem in the same market is not well-posed in the sense that it does not have any optimal solution. This negative result has then been extended to a general mean–downside-risk mode. Furthermore, for the model with a general convex risk measure, delicate analysis has been carried out to obtain a complete solution. By the similar way, we also analysis the fixed-target problem. At last, we have showed that in a single-period market, the mean-semivariance problem admits optimal solutions as long as it is feasible, and so does the mean-downside-risk problem under some condition. The results in this chapter suggest that there are strikingly differences between the single-period and continuous-time markets.

Chapter 6

Conclusion Remark

A half century ago, H. Markowitz pioneered the modern finance theory by his mean-variance portfolio selection model. Although his work is perhaps technically simple in today's view, his idea still inspires the work in finance.

Zhou and Li [70] explored Markowitz's work in a complete continuous-time financial market. In this thesis, I investigated the continuous-time portfolio selection problem in more details, including in complete markets and incomplete markets, with constraints and without constraints. I also went beyond Markowitz's mean-variance framework by studying the mean-risk portfolio selection problem, and showed some properties of the problem which are totally different from those in the single-period or multi-period markets.

Even within the mean-variance framework, the continuous-time model is also quite different from the discrete-time one. In Zhou and Li [70] as well as in Chapter 3 and Chapter 4 in this thesis, I find that when the parameters are deterministic, most elegant properties of the mean-variance model in single period, such as the efficient frontier, the two-fund theorem, *etc.*, are carried over to the continuous-time model. Contrary to common belief, when the parameters are deterministic, the mean-variance portfolio optimization problems in continuous-time market are not really complex. In fact, when there are some additional constraints, the sharper results can be obtained in the continuous-time market than those in the single-period market.

When the parameters are stochastic—there is no counterpart in the single-period—the portfolio optimization problem will becomes very complex, especially in the case when the market is incomplete and some constraints are present. In this thesis, I overcome the difficulty by dual method. By constructing dual problems, the optimal solutions can be uniquely determined by the optimal solutions of the corresponding dual problems.

Variance has been commonly taken as a measure of risk. Meanwhile there are a lot of researches on how to measure the risk of an investment, such as semivariance advised by Markowitz. In the mean-risk framework, only mean-variance was widely accepted in the discrete-time market. One reason is there are only a few analytical results obtained for mean-risk models other than the mean-variance one. In the continuous-time case, the situation is different. In Chapter 5, we studied the mean-risk problem. A surprising conclusion is that other than a trivial case, a mean-downside-risk portfolio selection problem, such as the mean-semivariance one, admits no optimal solutions. For a general mean-risk portfolio optimization problem, we also need conditions to ensure the existence of optimal solutions. These conditions found in Chapter 5 are sufficient and necessary. Furthermore, I obtain an optimal solution if the existence is ensured, and find an asymptotic optimal solution when optimal solutions do not exist.

Appendix A

Measurable Selection Theorems

Before we state the next lemma, we note that a set $A \subset [0,T] \times \Omega$ is said to be $\mathcal{F}_t$-progressive if the corresponding indicator function 1_A is $\mathcal{F}_t$-progressively measurable. The $\mathcal{F}_t$-progressive sets form a σ-field (see, e.g., [23, p.99]).

Lemma A.1 *Let $X \equiv \{X(t) : 0 \leq t \leq T\}$ be a given n-dimensional, $\mathcal{F}_t$-progressively measurable stochastic process. Assume that $S(t,\omega) := \{y \in \mathbf{R}^m : f(X(t,\omega), y) \leq 0\} \neq \varnothing$ for any $(t,\omega) \in [0,T] \times \Omega$, where $f : \mathbf{R}^n \times \mathbf{R}^m \to \mathbf{R}^k$ is jointly measurable in both variables and continuous in the second variable. Then the process $\alpha \equiv \{\alpha(t) : 0 \leq t \leq T\}$ defined as $\alpha(t,\omega) := \operatorname{argmin}_{y \in S(t,\omega)} |y|^2$ is also $\mathcal{F}_t$-progressively measurable.*

Proof: First of all, for each $(t,\omega) \in [0,T] \times \Omega$, $S(t,w)$ is a closed set, and the square function is strictly convex and coercive. Hence $\alpha(t,\omega)$ is well defined. Set $g(t,\omega) := |\alpha(t,\omega)|^2$. Then for any $x \in \mathbf{R}$,

$$\{(t,\omega) : g(t,\omega) < x\} = \cup_{v \in Q^n, |v|^2 < x} \{(t,\omega) : f(X(t,\omega), v) \leq 0\}.$$

This shows that g is $\mathcal{F}_t$-progressively measurable.

Denote $S_n(t,\omega) := S(t,\omega) \cap \{y \in \mathbf{R}^m : |y|^2 \leq g(t,\omega) + 1/n\}$, for $(t,\omega) \in [0,T] \times \Omega$, and $n = 1, 2, \cdots$. Fix n. For any open set $O \subset \mathbf{R}^m$, we have

$$\{(t,\omega) : S_n(t,\omega) \cap O \neq \varnothing\} = \cup_{v \in O \cap \mathbf{Q}^m} \{(t,\omega) : f(X(t,\omega), v) \leq 0,\ |v|^2 \leq g(t,\omega) + 1/n\},$$

which is therefore an $\mathcal{F}_t$-progressive set. This shows that $S_n(t,\omega)$ satisfies the condition required in the measurable selection theorem [32, p. 281, Theorem 8.3.ii]. Hence, there exists an $\mathcal{F}_t$-progressively measurable process α_n with $\alpha_n(t,\omega) \in S_n(t,\omega)$ almost surely on $[0,T] \times \Omega$. It is clear that $\alpha_n(t,\omega) \to \alpha(t,\omega)$, almost surely, as $n \to \infty$. Thus α is $\mathcal{F}_t$-progressively measurable. $\square$

Lemma A.2 *$Y \equiv \{Y(t) : 0 \le t \le T\}$ be a given n-dimensional $\mathcal{F}_t$-progressively measurable stochastic process. Assume that $F(t,w) = \mathrm{argmin}_{d \in C} h(Y(t,\omega), d) \ne \varnothing$ for any $(t,\omega) \in [0,T] \times \Omega$, where C is a closed subset of $\mathbf{R}^m$, $h : \mathbf{R}^n \times C \to \mathbf{R}^k$ is jointly measurable in both variables and continuous in the second variable on C. Then the process $d \equiv \{d(\cdot) : 0 \le t \le T\}$ defined as $d(t,\omega) \stackrel{\triangle}{=} \mathrm{argmin}_{d \in F(t,\omega)} |d|^2$ is also $\mathcal{F}_t$-progressively measurable.*

Proof: Define $g(t) = \min_{d \in C} h(Y(t), d)$, then

$$\{(t,\omega) : g(t,\omega) < x\} = \cup_{v \in Q^n \cap C} \{(t,\omega) : h(Y(t,\omega), v) < x\}.$$

This shows that g is $\mathcal{F}_t$-progressively measurable.

Define $X(t) = (Y(t), g(t))', f(X(t), d) = |h(Y(t), d) - g(t)| + \min_{y \in C} |d - y|^2$, then X and f satisfies the conditions in Lemma A.1, and $d(t,\omega) = \mathrm{argmin}_{f(X(t),d) \le 0} |d|^2$, by Lemma A.1, d is $\mathcal{F}_t$-progressively measurable. $\square$

Appendix B

Some Lemmas on Vector Optimization

We start by recalling properties of a pseudo matrix inverse [51]. Let a matrix $M \in \mathbf{R}^{m\times n}$ be given. Then there exists a unique matrix $M^\dagger \in \mathbf{R}^{n\times m}$, called the Moore–Penrose pseudo inverse of M such that

$$\begin{cases} MM^\dagger M = M, \quad M^\dagger M M^\dagger = M^\dagger, \\ (MM^\dagger)' = MM^\dagger, \quad (M^\dagger M)' = M^\dagger M. \end{cases} \tag{B.1}$$

Lemma B.1 *Let matrices L, M, and N be given with appropriate dimension. Then the matrix equation*

$$LXM = N \tag{B.2}$$

has a solution X if and only if

$$LL^\dagger N M^\dagger M = N. \tag{B.3}$$

Proof: See [1, Lemma 2.7]. $\square$

Lemma B.2 *Let $A \in \mathbf{R}^{m\times n}$ and $b \in \mathbf{R}^m$. Then the following optimization problem*

$$\min_{y\in\mathbf{R}^n} |Ay - b|^2 \tag{B.4}$$

admits optimal solutions.

Proof: Since (B.4) is an unconstrained convex optimization problem, the zero-gradient condition yields that y is optimal if and only if $A'Ay = A'b$. By Lemma B.1, the latter has a solution if and only if

$$(A'A)(A'A)^{\dagger}A'b = A'b. \tag{B.5}$$

Denote $M := A'[I - A(A'A)^{\dagger}A']$. Then $MM' = A'[I - A(A'A)^{\dagger}A'][I - A(A'A)^{\dagger}A']A = A'[I - A(A'A)^{\dagger}A']A = 0$ where we have repeatedly used (B.1). This implies that $M = 0$ and, hence, (B.5) is satisfied. $\square$

Lemma B.3 *Given $a \in \mathbf{R}^n$ and $A \in \mathbf{R}^{m\times n}$. If $a \notin \{A'u : u \in \mathbf{R}^m_+\}$, then there exists $v \in \mathbf{R}^n \setminus \{0\}$ such that $a'v = -1$ and $Av \geq 0$.*

Proof: By the assumption $a \neq 0$. Denote $M := \{w \in \mathbf{R}^n : a'w < 0\}, N := \{w \in \mathbf{R}^n : Aw \geq 0\}$, which are both nonempty convex cones. If $M \cap N = \emptyset$, then by the convex separation theorem, there exists $y \in \mathbf{R}^n \setminus \{0\}$ with $\sup_{w\in A} y'w \leq \inf_{w\in B} y'w$. This implies

$$y'w \leq 0 \;\; \forall w \text{ with } a'w < 0, \tag{B.6}$$

and

$$y'w \geq 0 \;\; \forall w \text{ with } Aw \geq 0. \tag{B.7}$$

It follows from (B.6) that there exists $k > 0$ such that $a = ky$. On the other hand, (B.7) together with Farkas' lemma (see, e.g., [3, p.58, Theorem 2.9.1]) imply that there exists $\pi \in \mathbf{R}^m_+$ such that $y = A'\pi$. So $a = ky = A'(k\pi) \in \{A'u : u \in \mathbf{R}^m_+\}$, leading to a contradiction. Hence $M \cap N \neq \emptyset$. The desired conclusion then follows immediately. $\square$

Appendix C

Some Properties of Convex Functions on R

In this appendix we present some properties of a convex function $f\colon \mathbf{R} \to \mathbf{R}$. Let such a convex function f be fixed, and $\partial f(x)$ be its subdifferential at $x \in \mathbf{R}$.

Proposition C.1 *For any interval $A \subset \mathbf{R}$, $\cup_{x\in A}\partial f(x)$ is a convex set (and hence is an interval).*

Proof: Suppose $y_1 \in \partial f(x_1), y_2 \in \partial f(x_2)$ where $x_1, x_2 \in A$ with $x_1 < x_2$ and $y_1 < y_2$. It suffices to show that for any $y_0 \in (y_1, y_2)$, there is $x_0 \in [x_1, x_2]$ such that $y_0 \in \partial f(x_0)$.

It follows from the convexity that $x_1 \in \text{argmin}_{x\in\mathbf{R}}\{f(x)-y_1x\}$ and $x_2 \in \text{argmin}_{x\in\mathbf{R}}\{f(x)-y_2x\}$. On the other hand, the continuity of f ensures that there exists $x_0 \in [x_1, x_2]$ so that $f(x_0) - y_0x_0 = \min_{x\in[x_1,x_2]}\{f(x) - y_0x\}$. However, for any $x \leq x_1$,

$$\begin{aligned}
f(x) - y_0x &= f(x) - y_1x + (y_1 - y_0)x \\
&\geq f(x_1) - y_1x_1 + (y_1 - y_0)x \\
&= f(x_1) - y_0x_1 + (y_1 - y_0)(x - x_1) \\
&\geq f(x_1) - y_0x_1 \\
&\geq f(x_0) - y_0x_0.
\end{aligned}$$

Similarly we can prove that $f(x) - y_0 x \geq f(x_0) - y_0 x_0$ for any $x \geq x_2$. Therefore $x_0 \in \mathrm{argmin}_{x\in\mathbf{R}}\{f(x) - y_0 x\}$, which implies that $y_0 \in \partial f(x_0)$. □

Proposition C.2 *If f is strictly convex at x_0, then*

$$(\cup_{x<x_0}\partial f(x)) \cap (\cup_{x>x_0}\partial f(x)) = \emptyset. \tag{C.1}$$

Proof: If the conclusion is not true, then there are $x_1 < x_0 < x_2$ so that $f'_-(x_2) \leq f'_+(x_1)$. Hence $f'_-(x_2) = f'_+(x_1)$ due to the non-decreasing property of the subdifferential of f. However, the convexity of f yields

$$f'_+(x_1) \leq f'_-(x_0) \leq f'_+(x_0) \leq f'_-(x_2).$$

Thus, all the above inequalities become equalities which, in turn, implies that f is not strictly convex at x_0. □

Define a set-valued function G: $\cup_{x\in\mathbf{R}}\partial f(x) \to 2^{\mathbf{R}}$

$$G(y) := \{x \in \mathbf{R} : y \in \partial f(x)\}, \quad \forall y \in \cup_{x\in\mathbf{R}}\partial f(x).$$

If f is strictly convex, then $G(y)$ is a singleton for each y. In general, we have

Proposition C.3 *For any $y \in \cup_{x\in\mathbf{R}}\partial f(x)$, $G(y)$ is a closed interval in* $\mathbf{R}$.

Proof: First we prove that $G(y)$ is an interval. For any $x_1 \in G(y), x_2 \in G(y)$ with $x_1 \leq x_2$, and any $x \in (x_1, x_2)$, we have $f'_-(x) \leq f'_-(x_2) \leq y \leq f'_+(x_1) \leq f'_+(x)$. This implies $y \in \partial f(x)$, or $x \in G(y)$.

To show the closedness of $G(y)$, take $x_n \in G(y)$ with $x_n \to x \in \mathbf{R}$. Since $y \in \partial f(x_n)$, we have $f(x') - f(x_n) \geq y(x' - x_n)\ \forall x' \in \mathbf{R}$. This yields $f(x') - f(x) \geq y(x' - x)\ \forall x' \in \mathbf{R}$, implying that $y \in \partial f(x)$ or $x \in G(y)$. □

Now, define the function g: $\cup_{x\in\mathbf{R}}\partial f(x) \to \mathbf{R}$ as

$$g(y) := \mathrm{argmin}_{x\in G(y)}|x|, \ \ \forall y \in \cup_{x\in\mathbf{R}}\partial f(x).$$

Thanks to Proposition C.3, g is well defined.

Proposition C.4 *The set* $\{y \in \cup_{x\in\mathbf{R}}\partial f(x) : G(y)$ *is not a singleton*$\}$ *is countable.*

Proof: Take any $y_1 < y_2$ such that $G(y_1)$ and $G(y_2)$ are not singletons. It follows from Proposition C.3 that both $\mathrm{int}(G(y_1))$ and $\mathrm{int}(G(y_2))$ are nonempty. Moreover, $\mathrm{int}(G(y_1)) \cap \mathrm{int}(G(y_2)) = \emptyset$. Indeed, if it is not true, then there exist $a < b$ such that $[a,b] \subset G(y_1) \cap G(y_2)$, leading to $f'_+(a) \geq y_2 > y_1 \geq f'_-(b)$ which is impossible. This proves the desired result. $\square$

Denote $\Gamma := \{y \in \cup_{x\in\mathbf{R}}\partial f(x) : G(y)$ is a singleton$\}$.

Proposition C.5 g *is increasing on* $\cup_{x\in\mathbf{R}}\partial f(x)$*, and continuous at every* $y \in \Gamma$*.*

Proof: For any $y_1, y_2 \in \cup_{x\in\mathbf{R}}\partial f(x)$ with $y_1 < y_2$, if $x_1 := g(y_1) > g(y_2) =: x_2$, then $y_1 \geq f'_-(x_1) \geq f'_+(x_2) \geq y_2$, which is a contradiction. So $g(y_1) \leq g(y_2)$.

To prove the continuity at points in Γ, fix $y_0 \in \Gamma$ and let $x_0 := g(y_0)$. Since g is an increasing function, $\bar{x} := \lim_{y\downarrow y_0} g(y) \geq g(y_0) = x_0$. If $\bar{x} > x_0$, then for any $\epsilon > 0$ and $y > y_0$, one has $g(y) > \bar{x} - \epsilon$. Hence $y \geq f'_-(g(y)) \geq f'_+(\bar{x} - \epsilon)$, which implies

$$y_0 \geq f'_+(x_1 - \epsilon) \;\; \forall \epsilon > 0. \tag{C.2}$$

Now, for any $x \in (x_0, x_1)$ and $y \in \partial f(x)$, we have

$$y_0 \leq f'_+(x_0) \leq y \leq f'_+(x) \leq y_0$$

where the last inequality is due to (C.2). The above argument leads to $\cup_{x\in(x_0,x_1)}\partial f(x) = \{y_0\}$; so $G(y_0) \supseteq (x_0, x_1)$ is not a singleton, which contradicts that fact that $y_0 \in \Gamma$. This proves the right continuity of g. Similarly, one can show the left continuity of g. $\square$

Corollary C.1 *If* f *is strictly convex, then* g *is increasing and continuous on* $\cup_{x\in\mathbf{R}}\partial f(x)$*.*

Proof: In view of Proposition C.5, it suffices to prove $\Gamma = \cup_{x\in\mathbf{R}}\partial f(x)$ or, equivalently, $G(y)$ is a singleton for any $y \in \cup_{x\in\mathbf{R}}\partial f(x)$.

Suppose $[x_1, x_2] \subset G(y)$, then $y \leq f'_+(x_1) \leq f'_-(x_2) \leq y$. Hence $f'_+(x_1) = f'_-(x_2) = y$ which implies that $\partial f(x) = \{y\}$ for all $x \in (x_1, x_2)$. Therefore $f(\cdot)$ is not strictly convex on (x_1, x_2). $\square$

Define $g(y) := f(y) - (\lambda - \mu\xi)y, \quad y \in \mathbf{R}$, and $h := \inf_{y \in \mathbf{R}} g(y)$. If Y^* is not almost surely a minimum point of $g(\cdot)$, then $P(h < g(Y^*)) > 0$. Hence there exists $\epsilon > 0$ such that $P(h < g(Y^*) - \epsilon) > 0$. Since $(h < g(Y^*) - \epsilon) = \cup_{r \in \mathbf{Q}}(h < r, g(Y^*) > r + \epsilon)$, there exists $r \in \mathbf{Q}$ so that $P(h < r, g(Y^*) > r + \epsilon) > 0$.

Set $A := (h < r, g(Y^*) > r + \epsilon)$ and

$$B_n := (\text{there exists } y \in (-n, n) \text{ such that } g(y) < r + \epsilon, |y - Y^*| < n), \quad n \in \mathbf{N}.$$

Since $B_n = \cup_{y \in \mathbf{Q} \cap (-n,n)}\{g(y) < r + \epsilon, |y - Y^*| < n\}$ due to the continuity of g, each B_n is an $\mathcal{F}$-measurable set. Now,

$$\cup_{n \in \mathbf{N}}(A \cap B_n) = \{\text{there exists } y \in \mathbf{R} \text{ such that } g(y) < r + \epsilon\} \cap A \supseteq A;$$

hence $P(A \cap B_n) > 0$ for some fixed $n \in \mathbf{N}$, thanks to $P(A) > 0$. On $A \cap B_n$, define $Y := \sup\{y \in (-n, n) : g(y) < r + \epsilon\}$. Construct

$$Y' := \begin{cases} Y, & \omega \in A \cap B_n, \\ Y^*, & \omega \notin A \cap B_n. \end{cases}$$

Assuming that $Y' \in L^q(\mathcal{F}, \mathbf{R})$ (which will be proved below), we have $g(Y) < g(Y^*)$ on $A \cap B_n$ whose probability is greater than zero. Hence $Eg(Y') < Eg(Y^*)$, leading to a contradiction.

It remains to prove that Y' is $\mathcal{F}$-measurable and $Y' \in L^q(\mathcal{F}, \mathbf{R})$. For the measurability, it suffices to show that Y is a measurable random variable on the measurable space $(\Omega_0, \mathcal{F}_0)$, where $\Omega_0 := A \cap B_n$ and $\mathcal{F}_0 = \{\Gamma \cap \Omega_0 : \Gamma \in \mathcal{F}\}$, or equivalently, $\{Y \leq x\} \in \Omega_0\ \forall x \in \mathbf{R}$. Indeed, $\{Y \leq x\} = \emptyset$ for $x \leq -n$, $\{Y \leq x\} = \Omega_0$ for $x \geq n$, and $\{Y \leq x\} = \Omega_0 \setminus \cup_{y \in \mathbf{Q} \cap (x,n)}\{g(y) < r + \epsilon\}$ for $-n < x < n$. This proves the measurability of Y on $(\Omega_0, \mathcal{F}_0)$. Finally, $|Y| \leq n$ on Ω_0 which yields $Y' \in L^q(\mathcal{F}, \mathbf{R})$. $\square$

Bibliography

[1] M. Ait Rami, J.M. Moore and X.Y. Zhou, *Indefinite stochastic linear quadratic control and generalized differential Riccati equation*, SIAM Journal on Control and Opimization, 45 (2000), pp. 1131-1143.

[2] L. Bachelier, *Thérie de la Speculation*, Annales Scientifiques de L'Éole Normale Supérieure, 3rd ser., 17 (1990), pp. 21-88.

[3] A.V. Balakrishnan, *Applied Functional Analysis*, Springer–Verlag, New York, 1976.

[4] A.B. Berkelaar, *Strategic Asset Alllocation and Asset Pricing*, Erasmus Universiteit Rotterdam, Research Series 250, 2002.

[5] T.R. Bielecki, S.R. Pliska, H. Jin and X.Y. Zhou, *Continuous-time mean–variance portfolio selection with bankruptcy prohibition*, Math. Finance, to appear.

[6] F. Black, M. Scholes, *The pricing of options and corporate liabilities*, Journal of Political Economy, 81 (1973), pp. 635-654.

[7] X. Cai, K.L. Teo, X. Yang and X.Y. Zhou, *Portfolio optimization under a minimax rule*, Management Science, 46 (2000), pp. 957-972.

[8] J.C. Cox and C.F. Huang, *Optimal consumption and portfolio policies when asset prices follow a diffusion process*, Journal Economic Theory, 49 (1989), pp. 33-83.

[9] J.C. Cox and C.F. Huang, *A variational problem arising in financial economics*, Journal of Math. Economic, 20 (1991), pp. 465-487.

[10] J. Cvitanic and I. Karatzas, *Convex duality in constrained portfolio optimization*, Annals of Applied Probability, 2 (1992), pp. 767-818.

[11] F. Delbaen, W. Schwchermayer, *The Variance-optimal martingale measure for continuous process*, Bernoulli 2 (1996), pp. 81-105.

[12] D. Duffie and H.R. Richardson, *Mean variance hedging in continuous time*, Annals of Applied Probability, 1 (1991), pp. 1-15.

[13] N. El Karoui, S. Peng and M.C. Quenez, *Backward stochastic differential equations in finance*, Math. Finance, 7 (1997), pp. 1-71.

[14] R.J. Elliott and P.E. Kopp, *Mathematics of Financial Markets*, Springer-Verlag, New York, 1999.

[15] E. Javier, *Mean-Semivariance Behavior: The D-CAPM*, working paper, 2002, IESE Business School.

[16] P.C. Fishburn. *Mean-risk analysis with risk associated with below-target returns*, American Economic Review, 67 (1977), pp.116-126.

[17] C. Gourieroux, J.P. Laurent, H. Pham, *Mean-variance hedging and numeraire*, Math. Finance, 8 (1998), pp. 179-200

[18] R.R. Grauer and N.H. Hakansson, *On the use of mean–variance and quadratic approximations in implementing dynamic investment strategies: A comparison of the returns and investment policies*, Management Science, 39 (1993), pp. 856-871.

[19] N.H. Hakansson, *Capital growth and the mean–variance approach to portfolio management*, Journal of Financial Quantitative Analysis, 6 (1971), pp. 517-557.

[20] Y. Hu and X.Y. Zhou, *Constrained stochastic LQ control with random coefficients, and application to mean-variance portfolio selection*, preprint IRMAR, 03-23, 2003.

[21] C.F. Huang and R.H. Litzenberger, *Foundations for Financial Economics*, Pearson Education POD, 1998.

[22] P. Jorion, *Value at Risk: The New Benchmark for Managing Financial Risk*, McGraw-Hill, New York, 2001.

[23] O. Kallenberg, *Foundations of Modern Probability*, Springer-Verlag Telos, 1997.

[24] I. Karatzas, J.P. Lehoczky and S.E. Shreve , *Optimal portfolio and consumption decisions for a "small investo" on a finite horizon*, SIAM Journal on Control and Optimization, 25 (1991), pp. 1557-1586.

[25] I. Karatzas, J.P. Lehoczky, S.E. Shreve and Gan-lin Xu, *Martingale and Duality Methods for Utility Maximization in An Incomplete Market*, SIAM Journal on Control and Optimization, 29 (1991), pp. 702-730.

[26] I. Karatzas and S.E. Shreve, *Methods of Mathematical Finance*, Springer-Verlag, New York, 1998.

[27] M. Kohlmann and X.Y. Zhou, *Relationship between backward stochastic differential equations and stochastic controls: A linear-quadratic approach*, SIAM Journal on Control and Optimization., 38 (2000), pp. 1392-1407.

[28] H. Konno and H. Yamazaki, *Mean-absolute deviation portfolio optimization model and its application to tokyo stock exchange*, Management Science, 37 (1991), pp. 519-531.

[29] R. Korn, *Optimal Portfolios*, World Scientific, Singapore, 1997.

[30] R. Korn and S. Trautmann, *Continuous-time portfolio optimization under terminal wealth constraints*, ZOR, 42 (1995), pp. 69-93.

[31] D. Kramkov and W. Schachermayer, *Necessary and sufficient conditions in the problem of optimal investment in incomplete markets*, preprint, 2001.

[32] C. Lamberto, *Optimization–Theory and Applications*, Springer-Verlag, 1985

[33] D. Li and W.L. Ng, *Optimal dynamic portfolio selection: Multiperiod mean–variance formulation*, Math. Finance, 10 (2000), pp. 387-406.

[34] X. Li, X.Y. Zhou and A.E.B. Lim, *Dynamic mean–variance portfolio selection with no-shorting constraints*, SIAM Journal on Control and Optimization, 40 (2001), pp. 1540-1555.

[35] A.E.B. Lim, *Quadratic hedging and mean-variance portfolio selection with random parameters in an incomplete market*, Mathematics of Operations Research, 29 (2004), pp. 132-161

[36] A.E.B. Lim and X.Y. Zhou, *Mean–variance portfolio selection with random parameters*, Mathematics of Operations Research, 27 (2002), pp. 101-120.

[37] J. Lintner, *The valuation of risk assets and the selection of risky investment in stock portfolios and capital budgets*, Review on Economic Statistics, 47 (1965), pp. 13-37.

[38] D.G. Luenberger, *Optimization by Vector Space Methods*, Wiley, New York, 1969.

[39] J. Ma, P. Protter and J. Yong, *Solving forward-backward stochastic differential equations explicitly – A four step scheme*, Probability Theory and Related Fields, 98 (1994), pp. 339-359.

[40] J. Ma and J. Yong, *Forward-Backward Stochastic Differential Equations and Their Applications*, Lecture Notes on Math., Vol. 1702, Springer-Verlag, New York, 1999.

[41] H. Markowitz, *Portfolio selection*, Journal of Finance, 7 (1952), pp. 77-91.

[42] H. Markowitz, *Portfolio Selection: Efficient Diversification of Investments*, John Wiley & Sons, New York, 1959.

[43] H. Markowitz, *Mean–Variance Analysis in Portfolio Choice and Capital Markets*, Frank J. Fabozzi Associates, New Hope, 1987.

[44] H. Markowitz, *Foundation of Portfolio Theory*. Nobel Lecture, Dec 1990

[45] R. Merton, *Optimum consumption and portfolio rules in a continuous time model*, Journal of Economic Theory, 3 (1971), pp. 373-413; Erratum 6 (1973), pp. 213-214.

[46] R. Merton, *Continuous-Time Finance*, Blackwell Publishers, pp. 213-214, 1992.

[47] M. Magill, M. Quinzil and M. Quinzii, *Theory of Incomplete Markets, Vol. 1*, MIT Press; Reprint edition, 2002

[48] J. Mossin, *Equilibrium in a capital asset market*, Econometrica, 34 (1966), pp. 768-783.

[49] D. Nawrocki, *A brief history of downside risk measures*, Journal of Investing, 8 (1999), pp. 9-25.

[50] W. Ogryczak and A. Ruszczynski, *Dual stochastic dominance and related mean-risk models*, SIAM Journal on Optimization, 13(2002), pp. 60-78

[51] R. Penrose, *A generalized inverse of matrices*, Proc. Cambridge Philos. Soc., 52 (1955), pp. 17-19.

[52] S.R. Pliska, *A discrete time stochastic decision model*, Advances in Filtering and Optimal Stochastic Control, edited by W.H. Fleming and L.G. Gorostiza, Lecture Notes in Control and Information Sciences, 42, Springer-Verlag, New York, 290-304, 1982.

[53] S.R. Pliska, *A stochastic calculus model of continuous trading: Optimal portfolios*, Mathematics of Operations Research, 11 (1986), pp. 371-384.

[54] S.R. Pliska, *Introduction to Mathematical Finance*, Blackwell Publishers, Oxford, 1997.

[55] H.R. Richardson, *A minimum variance result in continuous trading portfolio optimization*, Management Science, 35 (1989), pp. 1045-1055.

[56] R.T. Rockafellar, *Convex Analysis*, Princeton University Press, Princeton, 1970.

[57] P.A. Samuelson, *Lifetime portfolio selection by dynamic stochastic programming*, Review on Economic Statistics, 51 (1986), pp. 371-382.

[58] P.A. Samuelson, *Economics*, McGraw-Hill, New York, 1948.

[59] M. Schweizer, *Mean–variance hedging for general claims*, Annals of Applied Probability, 2 (1992), pp. 171-179.

[60] W. Schachermayer, *Optimal investment in incomplete markets when wealth may become negative*, Annals of Applied Probability, 11 (2001), pp. 694-734.

[61] W.F. Sharpe, *Capital asset prices: A theory of market equilibrium under conditions of risk*, Journal of Finance, 19 (1964), pp. 425-442.

[62] F.A. Sortino and R. van der Meer, *Downside Risk*, Journal of Portfolio Management, 17 (1991), pp. 27-31.

[63] M.C. Steinbach, *Markowitz revisited: Mean–variance models in financial portfolio analysis*, SIAM Review, 43 (2001), pp. 31-85.

[64] J. von Neumann, O. Morgenstrern, *Theorey of Games and Economic Behavior*, Princeton Unviersity Press, 1953.

[65] J.M. Xia and J.A. Yan, *A new look at the fundamental theorem of asset pricing*, Journal of Korean Math. Society, 35 (1998), pp. 659-673.

[66] J.A. Yan, *A Numeraire-Free and Original Probability Based Framework for Financial Markets*, Priprint, 2000.

[67] J. Yong and X.Y. Zhou, *Stochastic Controls: Hamiltonian Systems and HJB Equations*, Springer, New York, 1999.

[68] Y. Zhao and W.T. Ziemba, *Mean-variance versus expected utility in dynamic investment analysis*, Working Paper, Faculty of Commerce and Business Administration, University of British Columbia, 2000.

[69] X.Y. Zhou, *Markowitz's world in continuous time, and beyond,* Stochastic Modeling and Optimization, edited by D.D. Yao *et al.*, Springer, New York, pp. 279-310, 2003.

[70] X.Y. Zhou and D. Li, *Continuous time mean–variance portfolio selection: A stochastic LQ framework*, Applied Mathematics and Optimization , 42 (2000), pp. 19-33.

[71] X.Y. Zhou and G. Yin, *Markowitz's mean-variance portfolio selection with regime switching: A continuous-time model*, SIAM Journal on Control and Optimization, 42 (2003), pp. 1466-1482.