ON CAPITAL REQUIREMENTS AND OPTIMAL STRATEGIES TO ACHIEVE ACCEPTABILITY

by

Soumik Pal

Under the supervision of

Professor Ioannis Karatzas

Submitted in partial fulfillment of the
requirement for the degree
of Doctor of Philosophy
in the Graduate School of Arts and Science

COLUMBIA UNIVERSITY

2006

Abstract

This thesis is concerned with an agent in a financial market who wishes to trade between day $t = 0$ and day $t = \tau$. The agent starts with a specific objective in mind, which defines a class of *acceptable* (random) financial positions at the end of day τ. He wants to determine an initial amount and a self-financing trading strategy which will turn the discounted terminal value of his portfolio acceptable. By suitable choices of acceptable positions, this simple structure becomes flexible enough to encompass most of the common problems in mathematical finance. There are at least two distinct viewpoints of acceptability: one, which comes from the theory of *measures of risk*, stipulates 'proper' measures of risk as a real function evaluated on future random net worths. The set of acceptable net worths are then defined by their non-positive risks. The other is a generalization of *arbitrage opportunities*. It defines acceptable positions as what everybody finds as *good deals*, and defines prices which are *fair* for both the buyer and the seller, in the sense that it rules out such good deals for both of them. In this thesis, we propose results determining the minimum capital requirement and a self-financing trading strategy via which the aforementioned agent can lead his wealth towards acceptability. Such minimum capital requirement are generalizations of the classical superhedging prices, although the mathematical techniques we apply are novel. The trading strategies can also be interpreted as generalized hedging strategies, and we propose new theoretical and computational methods to evaluate them.

Acknowledgements

In my experience research advances like the running maximum of a Brownian motion. At any given moment, its going nowhere, but on a set of Lebesgue zero measure it moves continuously upwards, ultimately hitting every possible value. I owe such continuous upward movements of my research, culminating in this thesis, to several people.

First, my advisor Prof. Ioannis Karatzas, who is a model of what is takes to be among the best in the world. I have throughout tried to emulate him as a researcher, a scholar, and a teacher. His immense contribution to my knowledge is strewn all over this thesis.

Before I came to Columbia, I spent five wonderful years at the sanctuaries of the Indian Statistical Institute, Calcutta. I take this opportunity to thank Prof. B. V. Rao and Prof. A. K. Roy for displaying the magical world of Probability and Analysis for the first time.

I would also like to thank my defense committee: Prof. Karatzas, Prof. Peter Bank, Prof. Savas Dayanik, Prof. Daniel Rabinowitz, and Prof. Jan Vecer for their patient reading of this lengthy thesis. Prof. Bank and Prof. Vecer deserve special thanks for their friendship and also for some very helpful academic discussions.

In several ways, my friend Dr. Sourav Chatterjee has always been a torch for me. One, his extraordinary insights in Mathematics has lit so many dark corners of my understanding. Two, his blazing faith in academic life constantly inspires me. I also thank him for being a very good friend of mine for all these years.

Thanks are also due to Mr. Rik Sen, who, over countless lunches and dinners, and strolls through the villages of Manhattan, explained to me the nuances of the 'real' financial world.

Life at the Statistics department at Columbia University had been full of fun. It is rare to get a department where all the faculties and students are so friendly. I will dearly miss mutual procrastination in the company of such great friends like Mr. Jouni Kerman, Mr. Jinfeng Xu, and Mr. Yu Zheng.

This list would be incomplete if I do not mention this constantly changing, but forever youthful New York City. This whirlwind of humanity, woven in myriad colors, has kept me amused even at the darkest hours.

Finally, there are two persons to whom I do not have the temerity to thank. My parents, Mr. Saroj Pal and Mrs. Sujata Pal, who braved tremendous hardships with honesty, courage and perseverance, in the hope that their only son would have the education they never got. They will not understand anything in this thesis, but everything here owes it to them.

I dedicate this thesis to my parents.

Contents

Chapter 1

Introduction

Central to what we do in this thesis is the idea of *acceptable* financial positions. This is a theory which has generated a lot of attention in recent mathematical finance literature, partly because it makes a rigorous assessment of risks associated with random financial net worths, and partly because it hugely generalizes No-Arbitrage asset pricing and superhedging ideas in incomplete markets. It took several authors and the last ten years, if not more, to fully explore the possibilities of such an idea. In the next chapter we go over the history of the development of the notion of acceptability, briefly touching the key facts, ideas and main results. A summary of our own contribution to this field made in this thesis, and how they relate to existing ideas, can be found in the final subsection 2.7.

In Chapter 3, we are concerned with the problem of minimum capital requirement which will allow an investor, by carefully trading in stocks and the money market, to reach *acceptability.* Such capital requirements can be interpreted as either sellers price of a contingent claim when hedging allows *acceptable shortfall* (superhedging being a particular case), or as a generalization of No-Arbitrage valuation principle where *good-deals* have been ruled out for both the buyer and the seller of the contingent claim. The mathematical set-up assumes acceptable positions corresponding to arbitrary convex measure of risk, and continuous trading on stock price processes modeled by special semimartingales. As far as we know, they give the most general kinds of such results known in the literature. We end this chapter with an example of computing sellers price of an European call options where we allow shortfall with bounded pth moment.

In Chapter 4, we take up the problem of determining strategies to achieve acceptability. The current literature is almost silent here, in part because a complete solution would provide strategies for almost any kind of hedging / portfolio optimization problem. We attack this problem by first reducing its scope. We consider discrete time trading, and acceptable positions defined by finitely many scenario probabilities and floors. In other words, the set of acceptable positions is a convex set of random variable determined by whether their expectations under finitely many probability measures dominate a corresponding level or not. Such situations have been studied either on their own merit or as an approximation to the general case. Under very mild assumptions, we can then greatly reduce our search for a suitable strategy to a small, convenient class. We then show how a Monte-Carlo procedure can approximate a valid strategy to arbitrary degree of precision at the cost of computing power.

We end this thesis with directions for possible future research.

Note. An index of commonly used mathematical symbols and the corresponding equation numbers where they have been introduced can be found in the next page.

1.1 Index of notations

Commonly used mathematical symbols in this thesis:	Key to deciphering, e.g., equation number where introduced:
$\mathcal{A}_\rho$	(2.2), (2.8)
$V@R_\alpha$	(2.3)
S_t	(3.1)
Θ	(3.2)
$W_u^{x,\pi}$	(3.3)
G	(4.19)
Λ	(3.5)
$\widetilde{\Lambda}$	(Closed) Convex hull of Λ
Γ	(3.6)
$\mathcal{A}_0$	(3.7)
f	(**??**)
$\tilde{f}$	(3.14)
$\mathfrak{T}$	(3.31)
$\overline{\mathcal{A}}_0$	(3.30)
$\mathcal{Z}$	(3.33)
$d_p(X, \Pi)$	(3.35)
S_Π	(3.39)
$W(\xi)$	(4.4)
$\mathcal{S}_{m+1}$	(4.60)
$v_t(f_i)$	(4.7)

Chapter 2

A brief history of acceptability

2.1 Acceptability and coherent measures of risk

One of the first articles to define and study acceptability is the seminal paper [Artzner et al., 1999]. They provide a definition of *risks* and present and justify a unified framework for analysis, construction and implementation of *measures of risk*. As the authors point out, these measures of risks, named *coherent* measures, can be used as extra capital requirements, to regulate the risk assumed by market participants, traders, insurance underwriters, as well as to allocate existing capital. The idea is twofold: first to stipulate axioms which define acceptable future random net worths, and secondly, to define the measure of risk of an unacceptable position as the minimum extra capital which, invested in a 'pre-specified reference investment instrument', makes the future discounted value of the position acceptable. The axioms defining acceptability do not specify a unique measure of risk, instead, they characterize a large class of risk measures. The choice of precisely which measure to use from this class is left to additional economic considerations.

Their basic object of study are random variables on the set of *states of nature* at a future date, interpreted as possible future values of positions or portfolios currently held. A supervisor (e.g. regulator, exchange's clearing firm, or investment manager) decides on a subset of such future outcomes as *acceptable risks*. Mathematically, they choose a subset $\mathcal{A}$ of the set of all real functions, $\mathbf{L}^0$, on a finite set Ω, and call it the acceptance set. A measure of risk associated with $\mathcal{A}$ is a function $\rho_{\mathcal{A}} : \mathbf{L}^0 \to \mathbb{R}$, defined by

$$\rho_{\mathcal{A}}(X) \triangleq \inf\{m \mid m + X \in \mathcal{A}\}.$$

Remark. This definition differs slightly from the original in [Artzner et al., 1999] since we ignore the rate of return on the reference instrument (e.g. interest rate in the money market).

Conversely, for any function $\rho : \mathbf{L}^0 \to \mathbb{R}$, one can define a corresponding acceptance set by

$$\mathcal{A}_\rho \triangleq \left\{X \in \mathbf{L}^0 \mid \rho(X) \leq 0\right\}.$$

The crucial point made in this paper is that a *proper* measure of risk, ρ, should satisfy the following axioms:

1. *Translation invariance:* for all $X \in \mathbf{L}^0$ and all real a, we have $\rho(X + a) = \rho(X) - a$.

2. *Subadditivity:* for all X_1 and X_2 in $\mathbf{L}^0$, we have $\rho(X_1 + X_2) \leq \rho(X_1) + \rho(X_2)$.

3. *Positive homogeneity:* for all $\lambda \geq 0$ and all $X \in \mathbf{L}^0$, we have $\rho(\lambda X) = \lambda \rho(X)$.

4. *Monotonicity:* for all X and Y in $\mathbf{L}^0$ with $X \geq Y$, we have $\rho(X) \leq \rho(Y)$.

If a function ρ satisfies all the above axioms, it is called a coherent measure of risk. This definition induces some restrictions on the corresponding acceptance set, $\mathcal{A}_\rho$, most importantly that it has to be a convex cone. Moreover, we also have $\rho_{\mathcal{A}_\rho} = \rho$. Why these axioms are natural requirements for a measure of risk has been argued in [Artzner et al., 1999, Section 2.2], and we skip such details.

The other major achievement in that paper is the following representation theorem.

Theorem 2.1 ([Artzner et al., 1999], Proposition 4.1)**.** *A risk measure ρ is coherent if and only if there exists a family $\mathcal{P}$ of probability measures on the set of states of nature, such that*

$$\rho(X) = \sup \left\{ \mathrm{E}^P[-X] \mid P \in \mathcal{P} \right\}. \tag{2.1}$$

And thus, the corresponding acceptance set is given by

$$\mathcal{A}_\rho = \left\{ X \in \mathbf{L}^0 \mid \mathrm{E}^P[X] \geq 0, \ \ \forall P \in \mathcal{P} \right\}. \tag{2.2}$$

These probability measures P, coined 'scenarios', are defined on Ω, where the σ-algebra is the power set of Ω. The authors also note that such a representation theorem had already been proved in Proposition 2, Chapter 10 of [Huber, 1981], albeit in a different context.

In the rest of the paper, the authors consider three existing measures of risk: the margin system Standard Portfolio Analysis of Risk (SPAN), developed by the Chicago Mercantile Exchange, the margin rules of the Securities and Exchange Commission (SEC), which are used by the National Association of Securities Dealers (NASD), and the value-at-risk V@R. They show that while the first two measures of risk are coherent, the last one is not. V@R is a family of functions depending on a probability measure P we put on Ω. Given $\alpha \in (0, 1)$, the function $V@R_\alpha$, defined as

$$V@R_\alpha(X) \stackrel{\triangle}{=} -\inf \left\{ x \mid P[X \leq x] > \alpha \right\}, \tag{2.3}$$

fails the subadditivity property and hence does not reward diversification. They prescribe a possible remedy in the coherent measure TailV@R, or Tail Conditional Expectation (TCE), which is a one-parameter family of coherent risk measures defined by

$$\mathrm{TCE}_\alpha \triangleq \mathrm{E}^P[-X \mid X \leq -V@R_\alpha(X)].$$

2.2 A generalization of the NA principle

In a now well-known paper, [Carr et al., 2001], the authors use the notion of acceptability to present a new approach for positioning, pricing, and hedging in incomplete markets that bridges standard arbitrage pricing and expected utility maximization. Their starting point is the observation that an arbitrage is an *opportunity* that absolutely everyone would accept. It follows, from the continuity of preferences, that opportunities exist, with mild risks, which would be considered acceptable to all but the most risk-averse. They term such an opportunity that is agreeable to a wide variety of reasonable individuals to be an *acceptable opportunity.* Such a class of acceptable opportunities is specified by a finite set of probability measures and associated floors, in the sense that an investment is deemed acceptable if its expected pay-offs under each probability measure exceed the corresponding floor. In other words, we are given a pre-specified set of probability measures $P_i, 1 \leq i \leq m$, on the future states of the world, and associated floors $f_i,\ 1 \leq i \leq m$. If X denotes the final position of an investment strategy, the strategy will be acceptable if

$$\mathrm{E}^{P_i}[X] \geq f_i, \quad i = 1, \ldots, m. \tag{2.4}$$

The paper requires the constants $\{f_i\}$ to be non-positive. For example, if Ω is finite, a suitable choice of P_i (Dirac mass at sample point i) would yield arbitrage opportunities as the only acceptable ones.

The authors then proceed to extend the fundamental theorems of asset pricing. They examine the implications of the liquid assets being priced so that there are no strictly acceptable opportunities among them. Correspondingly, the concept of hedging is reformulated to one of attaining acceptable residual risks. We do not discuss further details of this paper, since very soon we shall see a more unified picture coming up. For the moment, observe the similarities of the two notions of acceptability (2.2) and (2.4), although they come from seemingly disparate considerations.

2.3 Coherent risk measures and NGD pricing

The authors of [Carr et al., 2001] were not the first one to consider bridges between No Arbitrage theory and utility maximization principle. The theory of *no-good-deal pricing* (NGD), as a pricing technique based on the absence of attractive investment opportunities in equilibrium, was introduced before them in [Černy and Hodges, 2001]. The term no-good-deal is borrowed from an earlier paper with similar objectives, [Cochrane and Saá-Requejo, 2000], where good-deals were defined by high sharp-ratio of returns. The first paper which fully establishes the link between coherent risk measures and the no-good-deal pricing theory is [Jaschke and Küchler, 2001]. One of their key results is that coherent risk measures are essentially equivalent to good-deal bounds. The key idea is to observe that the set of *desirable claims* has a one-to-one correspondence with the set of acceptable risks, in the sense of [Artzner et al., 1999], and thus gives rise to a coherent measure of risk ρ. Then they make the following crucial observation: let M be a set of *cash streams available in the market* (i.e., cash streams that can be generated by trading without endowments), then one can define a new risk measure

$$\tilde{\rho}(X) \triangleq \inf_{Y \in M} \rho(X+Y), \tag{2.5}$$

which is again coherent if M is a cone. The good-deal bounds, $\underline{\pi}$ and $\overline{\pi}$, which are generalized valuation bounds, then correspond to the coherent risk measure $\tilde{\rho}$ by $\tilde{\rho}(X) = \overline{\pi}(-X) = -\underline{\pi}(X)$. Such price intervals $[\underline{\pi}, \overline{\pi}]$ are consistent with the absence of good-deals in the market. Correspondingly, as we mention in the last section, the idea of hedging has been expanded to one where the residual risk is acceptable. Relations between measures of risk and No Good Deals are further extended by Staum in [Staum, 2004], where he proves fundamental theorem of asset pricing for good deal bounds in incomplete markets.

2.4 Convex measures of risk

A significant extension, encompassing all the previous ideas we discussed in this section, was made by introducing convex measures of risk in [Föllmer and Schied, 2002]. They consider the same set-up as in [Artzner et al., 1999]. For example, future discounted net worth from any financial position are represented by the vector space $\mathbf{L}^0$ of functions from Ω to $\mathbb{R}$. A quantitative measure of risk is real function ρ, defined on $\mathbf{L}^0$, which can be interpreted as a *margin requirement*, i.e., the minimal amount of capital, which if added to the position at the beginning of the given period and invested into a risk-free asset, makes the discounted position X acceptable. What the authors argue is that the positive homogeneity of the coherent risk measure is an undue requirement, because the risk of a position might increase in a non-linear way with the size of the position. For example, an additional liquidity risk may arise if a position is multiplied by a large factor. They suggest to relax the conditions of positive homogeneity and of subadditivity and to require the weaker property of convexity:

$$\rho(\lambda X + (1-\lambda)Y) \leq \lambda\rho(X) + (1-\lambda)\rho(Y), \quad \text{for any } \lambda \in [0,1],\ X, Y \in \mathbf{L}^0. \quad (2.6)$$

As with subadditivity, convexity means that diversification does not increase risk. A measure of risk is called a *convex measure of risk* if it satisfies (2.6) in addition to sharing the properties of Translation invariance and Monotonicity of the earlier coherent measures.

The authors of [Föllmer and Schied, 2002] then prove a representation theorem, similar in spirit to 2.1, which show that any convex measure of risk on a finite Ω is of the form

$$\rho(X) = \sup_{P \in \mathcal{P}} \left(\mathrm{E}^P[-X] - \alpha(P) \right). \quad (2.7)$$

Here, as before, the set $\mathcal{P}$ is the set of all probability measures on Ω. The function $\alpha(\cdot)$ is a certain *penalty function* on $\mathcal{P}$. When α is throughout zero, we get back the coherent measures. Representation (2.7) was independently proved by David Heath in [Heath, 2000]. As before, a convex measure of risk defines an associated acceptance set given by

$$\mathcal{A}_\rho = \left\{ X \in \mathbf{L}^0 \mid \rho(X) \leq 0 \right\} = \left\{ X \in \mathbf{L}^0 \mid \mathrm{E}^P[X] \geq -\alpha(P) \right\}. \quad (2.8)$$

Conversely, a given set of acceptable positions $\mathcal{A}$ defines a convex measure of risk via

$$\rho_{\mathcal{A}}(X) = \inf \{ m \in \mathbb{R} \mid m + X \in \mathcal{A} \} \quad (2.9)$$

provided $\mathcal{A}$ satisfies certain axioms.

Till now, all the measures of risk that we have discussed are *model-free* in the sense that no *base* probability has been assumed on the space Ω. In order to extend (2.7) to arbitrary Ω, we need to consider a base probability P on $(\Omega, \mathcal{F})$, where $\mathcal{F}$ is a suitable σ-algebra on Ω. Let $\mathbf{L}^\infty(P)$ denote the usual Banach space of P-essentially bounded real functions on Ω. Also, let $\mathbf{L}^1(P)$ denote the P-integrable functions on Ω. We have the following theorem.

Theorem 2.2 ([Föllmer and Schied, 2002] Theorem 6)**.** *Supose ρ is a convex measure of risk defined on $\mathbf{L}^\infty(P)$. Let $\mathcal{P}$ be the set of probability measures $Q \ll P$. Then the following properties are equivalent.*

1. *There is a penalty function $\alpha : \mathcal{P} \to (-\infty, \infty]$ such that*

$$\rho(X) = \sup_{Q \ll P} \left(\mathrm{E}^Q[-X] - \alpha(Q) \right), \quad \textit{for all } X \in \mathbf{L}^\infty(P). \tag{2.10}$$

2. *The acceptance set $\mathcal{A}_\rho$ is closed with respect to the weak* topology on $\mathbf{L}^\infty(P)$.*

3. *Fatou property: if the sequence $(X_n)_{n\in\mathbb{N}} \subseteq \mathbf{L}^\infty(P)$ is uniformly bounded, and X_n converges to some X in probability P, then $\rho(X) \leq \liminf_n \rho(X_n)$.*

4. *Continuity from above: if the sequence $(X_n) \subseteq \mathbf{L}^\infty(P)$ decreases to $X \in \mathbf{L}^\infty(P)$, then $\rho(X_n) \to \rho(X)$.*

We should mention here that there is no uniqueness property of the penalty function in the representation (2.10).

Some of the most useful examples of convex measures of risk come from considering shortfall risk which is introduced in [Föllmer and Schied, 2002, Section3]. As we shall see later, these are useful in constructing efficient hedging of contingent claim. Suppose that $l : \mathbb{R} \to \mathbb{R}$ is an increasing convex loss function which is not identically constant. For a position $X \in \mathbf{L}^\infty(P)$ introduce the expected loss $\mathrm{E}^P[l(-X)]$. If l vanishes on $(-\infty, 0]$, then $\mathrm{E}^P[l(-X)] = \mathrm{E}^P[l(X^-)]$ may be viewed as a quantitative assessment of the shortfall risk.

Let x_0 be an interior point in the range of l. A position $X \in \mathbf{L}^\infty(P)$ will be called acceptable if the expected loss is bounded by x_0. Thus we consider the class

$$\mathcal{A} \triangleq \left\{X \in \mathbf{L}^\infty(P) \mid \mathrm{E}^P[l(-X)] \leq x_0\right\} \tag{2.11}$$

of acceptable positions. This induces a measure of risk $\rho_\mathcal{A}(X)$ by relation (2.9). It is straightforward to verify that $\rho_\mathcal{A}$ satisfies all the axioms in order to be called a convex measure of risk. Additionally, since l is continuous as a finite-valued convex function on $\mathbb{R}$, the Fatou property of Theorem 2.2 is also satisfied. Thus $\rho_\mathcal{A}$ possesses a representation of the form (2.10), where the corresponding penalty function α_0 can be expressed in terms of the Fenchel-Legendre transform of l, i.e.,

$$l^*(z) \triangleq \sup_{x\in\mathbb{R}} \left(zx - l(x)\right).$$

Theorem 2.3 ([Föllmer and Schied, 2002] Theorem 10)**.** *Suppose $\mathcal{A}$ is the acceptance set given by* (2.11)*. Then for $Q \ll P$, a penalty function of $\rho_\mathcal{A}$ is given by*

$$\alpha_0(Q) = \inf_{\lambda>0} \frac{1}{\lambda}\left(x_0 + \mathrm{E}^P\left[l^*\left(\lambda\frac{\mathrm{d}Q}{\mathrm{d}P}\right)\right]\right).$$

Example 2.1. For example, if we take $l(x) = e^x$ and let $x_0 = 1$, then $\rho(X) = \log \mathrm{E}^P[\exp(-X)]$. In that case, the penalty function is given by the well-known variational formula of the relative entropy, namely $\alpha_0(Q) = H(Q|P)$, where the relative entropy of Q with respect to P is defined as

$$H(Q|P) \triangleq \begin{cases} \int \mathrm{d}Q/\mathrm{d}P \log\left(\mathrm{d}Q/\mathrm{d}P\right)\mathrm{d}P, & \text{if } Q \ll P,\\ +\infty & \text{otherwise.}\end{cases}$$

Example 2.2. Another commonplace example is the case when, for some $p > 1$,

$$l(x) = \begin{cases} \frac{1}{p}x^p & \text{if } x \geq 0,\\ 0 & \text{otherwise,}\end{cases} \tag{2.12}$$

Then it turns out that $\alpha_0(Q) = (px_0)^{1/p} \cdot \mathrm{E}^P\left[(\mathrm{d}Q/\mathrm{d}P)^q\right]^{1/q}$, where $q = p/(p-1)$.

2.5 Measures of risk in a financial market

2.5.1 Definition and motivation

The financial implications of this thesis heavily relies on the idea of *measure of risk in a financial market*, a phrase introduced in [Föllmer and Schied, 2002]. The idea itself is older, and we have already seen it in (2.5). Several authors have recently contributed to the development of this theory. We shall soon discuss [Barrieu and El Karoui, 2005a] and [Barrieu and El Karoui, 2005b], who establish these risk measures as special cases of *inf-convolution* of risk measures. It is only by considering such risk measures that we begin to see the connections between the theory of coherent and convex measures (Sections 2.1 and 2.4) on the one hand, and the generalizations of pricing and hedging theory to incomplete markets (Sections 2.2 and 2.3) on the other hand. In this section we discuss the discrete-time case. We shall use a similar mathematical set-up in Section 4. The ideas pass smoothly to continuous time, except for the regularity conditions, which we discuss in Section 3. Here we provide a non-technial expositions following [Föllmer and Schied, 2002, Section 4].

For simplicity, we consider a market where there is one *bank account* and one stock which is traded at $(T+1)$ time points $t = 0, 1, \ldots, T$. We denote the discounted price process of one share of the stock by $(X_0, X_1, \ldots, X_T)$. Consider a self-financing trading strategy which dictates holding ξ_t number of shares of the stock during the trading period $[t, t+1)$, and

$$V_t = V_0 + \sum_{k=1}^{t} \xi_k (X_k - X_{k-1})$$

is the associated value process for an initial endowment V_0.

Now consider a set of acceptable positions $\mathcal{A}$. In other words, a random financial position Z will be of acceptable risk if $Z \in \mathcal{A}$. More generally, we can think of Z to be acceptable, if the 'risky part' of Z can be hedged at no additional cost, i.e., we can find a suitable hedging portfolio ξ such that

$$Z + \sum_{k=1}^{T} \xi_k (X_k - X_{k-1}) \in \mathcal{A}. \tag{2.13}$$

If we think of Z as being the value of a large portfolio, we should avoid running into the regime of illiquidity when hedging Z. For instance, one may want to impose individual limits on the number of shares held of the risky asset. In this case, we can allow ξ_k to take values in some intervals $[a_k, b_k]$. Such constraints on the hedging portfolio were first suggested by Cvitanic and Karatzas in [Cvitanic and Karatzas, 1992] and [Cvitanic and Karatzas, 1993]. Now, (2.13) generates a larger class of acceptable positions

$$\tilde{\mathcal{A}} \triangleq \left\{ Z \;\middle|\; \exists\, \xi \text{ such that } Z + \sum_{k=1}^{T} \xi_k (X_k - X_{k-1}) \in \mathcal{A} \right\}. \tag{2.14}$$

Define the *market measure of risk corresponding to* $\mathcal{A}$ by

$$\tilde{\rho}(Z) = \inf \left\{ m \mid m + Z \in \tilde{\mathcal{A}} \right\}. \tag{2.15}$$

In particular, if the to start with the set $\mathcal{A}$ is the set of acceptable positions associated with a convex measure of risk ρ, the new measure of risk $\tilde{\rho}$ is again convex. Moreover, $\tilde{\rho}$ is given by

$$\begin{aligned}
\tilde{\rho}(Z) &= \inf \left\{ m \mid m + Z \in \tilde{\mathcal{A}} \right\} \\
&= \inf \left\{ m \;\middle|\; \exists\, \xi \text{ such that } m + Z + \sum_{k=1}^{T} \xi_k (X_k - X_{k-1}) \in \mathcal{A} \right\} \\
&= \inf_{\xi} \inf \left\{ m \;\middle|\; m + Z + \sum_{k=1}^{T} \xi_k (X_k - X_{k-1}) \in \mathcal{A} \right\} \\
&= \inf_{\xi} \rho \left(Z + \sum_{k=1}^{T} \xi_k (X_k - X_{k-1}) \right), \qquad \text{by (2.9).}
\end{aligned} \tag{2.16}$$

The expression in (2.16) is a special case of (2.5), where the measure of risk is convex and the available cash flow is generated by trading with zero capital. Thus, suppose, in the set-up of Section 2.3, good deals are defined by a convex measure of risk ρ, i.e., X is a *good deal* if $\rho(X) \leq 0$. Then the no-good-deals valuation bounds for the claim Z is given by the interval $[-\tilde{\rho}(Z), \tilde{\rho}(-Z)]$. In Chapter 3 we shall see how to evaluate $\tilde{\rho}(Z)$ for arbitrary risk measure ρ and any pay-off Z. The process ξ which attains the infimum in (2.16) can hence be thought as a hedging strategy. In Chapter 4 we shall discuss methods to compute such strategies.

We can now specialize the previous paragraph with the shortfall risk measures discussed in Theorem 2.3. These are closely related to the problem of *efficient hedging* or the problem of utility maximization; see, e.g., [Föllmer and Leukert, 2000] and [Schachermayer, 2000]. We briefly describe the problem. In a complete financial market a given contingent claim can be replicated by a self-financing trading strategy, and the cost of the replication defines the price of the claim. In incomplete markets, it is not possible to replicate every contingent claim perfectly. One can still stay on the safe side by a *super-replicating* or *superhedging* strategy. See [El Karoui and Quenez, 1995] and [Karatzas, 1996]. But from a practical viewpoint the price of super-replication is often prohibitively high. Also, it does not reflect the reality since it models an extreme risk-averseness. Investors are often inclined to take up opportunities of making profit together with small or controlled risk of a loss. Several authors have come up with ideas of *partial* hedging to redress this shortcoming. For example, in [Föllmer and Leukert, 1999], the authors discuss the possibility of super-replicating not with probability one, but with a preset high probability. In a sense, this is a dynamic version of the value-at-risk or V@R.

In [Föllmer and Leukert, 2000] the authors define *efficient hedging* by using shortfall risks in terms of a convex loss function l. The convexity of l correspond to risk aversion. The shortfall risk is the expectation of the loss function applied to the shortfall. Now, there can be two objectives of a hedging portfolio, one, to minimize the shortfall risk, given some capital constraint. Alternatively, one can prescribe a bound on the shortfall risk and minimize the cost. These efficient hedges allow investors to interpolate in a systematic way between the extremes of a perfect hedge and no hedge depending on the prescribed level of shortfall risk.

The groundwork for bringing in convex measures of risk in such problems has already been done in Theorem 2.3. We consider a convex loss function l which vanishes on $(-\infty, 0]$. As we have already mentioned, for such a loss function $l(-X) = l(X^-)$, where $X^- = \max(0, -X)$. We consider the set of acceptable financial positions $\mathcal{A}$ given by (2.11) for some suitable x_0. Theorem 2.3 describes the corresponding convex risk measure. For a contingent claim Z, consider the case when there exists a trading strategy ξ and a real m such that $m - Z + \sum_{k=0}^{T} \xi_k(X_k - X_{k-1}) \in \mathcal{A}$. By the definition of $\mathcal{A}$, this would imply

$$\mathrm{E}\left[l\left(-Z + m + \sum_{k=0}^{T} \xi_k(X_k - X_{k-1})\right)^-\right] = \mathrm{E}\left[l\left(Z - m - \sum_{k=0}^{T} \xi_k(X_k - X_{k-1})\right)^+\right] \leq x_0. \quad (2.17)$$

In other words, if one starts with a capital of m and follows the trading strategy ξ, then at the end of time T, the shortfall risk will be bounded above x_0. By (2.15), the minimum m for which there is a ξ satisfying (2.17) is given by $\tilde{\rho}(-Z)$, which one can also interpret as the price (upper or sellers price, compare with the good-deal bounds) of such an efficient hedge. The hedging strategy would be as before the process ξ which achieves the infimum in (2.16).

2.5.2 Examples of risk measures in a market

We can now give several examples to illuminate the ideas of the previous subsection. Also as in the last subsection, here too we use the discrete time framework, just for simplicity of technicalities. All the ideas in this section can be generalized smoothly to the continuous time framework, which we shall need anyway in Chapter 3. In all these examples we consider Ω to be the set of all possible states of the universe. A base probability measure P on some suitable σ-algebra of Ω represents the randomness in the outcomes. Any financial position is a random variable defined on Ω.

Example 2.3 (Superhedging). Consider a contingent claim Y, and define the set of acceptable positions by

$$\begin{aligned}\mathcal{A} \triangleq \{X \in \mathbf{L}^\infty(P) \mid \quad P(X \geq 0) = 1\} \\ = \left\{X \in \mathbf{L}^\infty(P) \mid \quad \mathrm{E}^Q[X] \geq 0, \quad \forall\, Q \ll P\right\}. \quad (2.18)\end{aligned}$$

It is immediate from (2.8) that $\mathcal{A}$ is derived from a convex measure of risk

$$\rho(X) = \sup_{Q \ll P} \left(\mathrm{E}^Q[-X] - \alpha(Q)\right), \quad \alpha(Q) \equiv 0.$$

If we now consider the market measure of risk corresponding to $\mathcal{A}$, defined in (2.15), we get

$$\tilde{\rho}(-Y) = \inf\left\{m \,\middle|\, \exists\, \xi \text{ such that } P\left[m + \sum_{k=1}^{T} \xi_k (X_k - X_{k-1}) \geq Y\right] = 1\right\}.$$

In other words, $\tilde{\rho}(-Y)$ is the no-arbitrage upper superhedging price for the contingent claim Y, and the corresponding strategy ξ (if it exists) is the superhedging strategy.

Example 2.4 (Hedging with shortfall risk). As before we consider the problem of hedging a contingent claim Y, but with controlled shortfall. Our choice of loss function for the purpose of this example is the pth moment i.e. $l(x) = (x^+)^p/p$. We have already discussed the risk measure arising out of this loss function, and in the line following (2.12) we describe the penalty function $\alpha_0(Q) = (px_0)^{1/p} \cdot \mathrm{E}^P[(\mathrm{d}Q/\mathrm{d}P)^q]^{1/q}$ for some pre-specified *tolerance level* $x_0 \in \mathbb{R}$. In other words, we define the set of acceptable positions by

$$\mathcal{A} \triangleq \left\{ X \in \mathbf{L}^\infty(P) \;\middle|\; \mathrm{E}[X^+]^p \leq px_0 \right\} = \left\{ X \in \mathbf{L}^\infty(P) \;\middle|\; \mathrm{E}^Q[X] \geq \alpha_0(Q) \right\}. \quad (2.19)$$

Just as in the previous example, we consider the market measure of risk of $-Y$, given by

$$\tilde{\rho}(-Y) = \inf \left\{ m \;\middle|\; \exists\, \xi \text{ such that } \mathrm{E}^P \left[Y - m - \sum_{k=1}^{T} \xi_k (X_k - X_{k-1}) \right]^+ \leq x_0 \right\}.$$

Thus $\tilde{\rho}(-Y)$ is the minimum amount of capital needed to have a shortfall risk bounded above by x_0.

Example 2.5 (Model uncertainty). Here we consider a different example of acceptability. Suppose we are uncertain about modeling the random outcomes of the financial market by P. Instead we have a collection of possible models $\{Q_1, Q_2, \ldots, Q_m\}$ all of which are probability measures on Ω. It makes sense to have each of these measures absolutely continuous to one another and absolutely continuous with respect to P. Otherwise, there will be events which are improbable under one, while probable under some of the others. This reflects an extreme level of uncertainty in modeling, and can be ruled out in many situations. Suppose the objective of an investor in this market is to generate a portfolio whose expected value will dominate another pay-off Y at the terminal time. But computing expectation depends on the choice of the probability measure. Hence a robust approach will be given by the following. Define the set of acceptable positions by

$$\mathcal{A} \triangleq \left\{ Z \in \mathbf{L}^\infty(P) \mid \mathrm{E}^{Q_i}[Z] \geq 0, \quad \text{for all } i = 1, 2, \ldots, m \right\}.$$

Hence the market measure of risk $\mathcal{A}$ of $-Y$, i.e., $\tilde{\rho}(-Y)$ is given by

$$\inf\left\{m \;\middle|\; \exists\, \xi \text{ such that } \mathrm{E}^{Q_i}\left[m + \sum_{k=1}^{T} \xi_k (X_k - X_{k-1})\right] \geq \mathrm{E}^{Q_i}[Y],\ 1 \leq i \leq m\right\}.$$

In other words $\tilde{\rho}(-Y)$ is the minimum capital required to generate, by trading, an expected pay-off more than that of Y.

2.6 Inf-convolution of risk measures

The measures of risk in a financial market is an example of risk transfer between agents (one of them consisting of the financial market). The right framework to develop this is through inf-convolution of risk measures. Although, for the purpose of this thesis, we do not need the full details of the theory, this might be a good place to introduce this important method of generating new risk measures from pre-existing ones. Readers interested in exploring other possibilities are referred to the excellent expository articles [Barrieu and El Karoui, 2005a] and [Barrieu and El Karoui, 2005b].

We start with the space $\mathbf{L}^\infty(P)$, which is the Banach space of all P-essentially bounded functions from Ω to $\mathbb{R}$. All the measures of risk we discuss in this subsection will be defined on this linear space. The *inf-convolution* of two convex functionals ϕ_A and ϕ_B on $\mathbf{L}^\infty(P)$ is defined as

$$\phi_A \square \phi_B(X) \triangleq \inf_{H \in \mathbf{L}^\infty(P)} \left\{\phi_A(X - H) + \phi_B(H)\right\}.$$

This is the functional extension of the classical inf-convolution operator acting on real convex functions $f \square g(x) := \inf_y \{f(x-y) + g(y)\}$.

Example 2.6. For example, suppose $\phi_A(X) = \mathrm{E}^{P_A}[-X]$, for some probability measure P_A on Ω, such that P_A is absolutely continuous with respect to P. And let ϕ_B

be equal to a convex measure of risk $\rho(X) = \sup_{Q \ll P}[\mathrm{E}^Q(-X) - \alpha(Q)]$. By using convex duality, it can be easily verified that the resulting inf-convolution will be given by the following convex measure of risk:

$$\phi_A \square \phi_B(X) = \mathrm{E}^{P_A}(-X) - \alpha(P_A).$$

Example 2.7. Another important class of examples is when ϕ_A is a convex measure of risk ρ, and ϕ_B is the *convex indicator* of a convex set B, i.e.,

$$\phi_B(X) = \begin{cases} 0, & \text{if } X \in B, \\ \infty, & \text{otherwise.} \end{cases}$$

It is straightforward to see that

$$\phi_A \square \phi_B(X) = \inf_{H \in \mathbf{L}^\infty(P)} \{\rho(X - H) + \phi_B(H)\} = \inf_{H \in B} \rho(X - H). \tag{2.20}$$

A quick comparison with expression (2.5) or, the more specialized version, (2.16), proves that measures of risk in a financial market are special cases of (2.20). In [Barrieu and El Karoui, 2005a], the authors refer to such reduction of risks, as in (2.20), as *transfer of risk*.

We end this subsection with the following important representation theorem of the inf-convolution of two convex measures of risk. This is a special case of Theorem 3.1 in [Barrieu and El Karoui, 2005a]. The central idea is that the Legendre-Fenchel transform of inf-convolutions act additively on each component in the following sense.

Theorem 2.4. *Let ρ_A, and ρ_B be two convex measures of risk on $\mathbf{L}^\infty(P)$ with corresponding penalty functions α_A and α_B. Let $\rho_A \square \rho_B$ denote their inf-convolution, and assume that $\rho_A \square \rho_B(0) > -\infty$. Then $\rho_A \square \rho_B$ is another convex measure of risk with associated penalty function $\alpha(Q) = \alpha_A(Q) + \alpha_B(Q), \quad \forall\ Q \in \mathcal{P}$, where $\mathcal{P}$ is the class of all probability measures on Ω absolutely continuous with respect to P.*

2.7 Summary of the thesis

This thesis is in two parts: in the first part, we are concerned with computing market measures of risk, defined in (2.15), when the set of acceptable positions $\mathcal{A}$ is defined through some convex measure of risk ρ, as in (2.8). Such attempts have been made by other authors in simpler set-ups, notably in [Larsen et al., 2004] who consider the case when ρ is determined by finitely many scenarios, and also [Cherny, 2005], who considers ρ to be a coherent measure of risk. Our propositions are the most general results of such kinds, in the sense that we assume continuous-time trading, semimartingale price processes, and arbitrary convex risk measures ρ. Such generalisations are essential as we have exhibited in examples 2.3, and 2.4 in Subsection 2.5.2. Please see Section 2.5 for a review of the wide applicability of market measures of risk. The mathematical hurdles one faces in the generality that we consider are considerable, and we employ theories such as that of *uniformly convex Banach spaces* which is uncommon in the literature. This has been done in Chapter 3, and the main results are stated in Proposition 3.4 and Proposition 3.5.

The second part of this thesis is focussed on determining the process ξ which achieves the infimum in (2.16). The current literature is almost silent on such problems, in part because a complete solution will determine hedging portfolios for almost all major efficient hedging problems, something which looks far-fetched. Here we resort to a computational approach in discrete time. Chapter 4 is devoted to this procedure. We consider again a market measure of risk induced by a convex measure ρ, described by finitely many scenarios and floors. In Proposition 4.1 we reduce the search of optimal ξ to a much smaller, more convenient class. The technique we use is a novel application of the Neyman-Pearson lemma. We then describe how this reduction allows us to do a simple, intuitive Monte-Carlo procedure to determine an approximately optimal ξ. The precision can be chosen arbitrarily at the cost of computing power. We should mention here that the application of Neyman-Pearson lemma to portfolio optimization problems by itself is not new. For example a beautiful application of this theory to quantile hedging can be found in [Föllmer and Leukert, 1999]. However, our approach and the mode of application of the Neyman-Pearson theory is different and new.

Each of Chapters 3 and 4 ends with examples applying our results, followed by an appendix which details the more involved proofs and short introductions to some of the non-standard mathematical theories we employ.

We conclude this thesis with a short discussion about unresolved problems and possible future directions of work.

Chapter 3

Capital requirement to achieve acceptability

3.1 Introduction

Consider an agent who trades during a finite time-interval $[0, T]$ in a market that offers finitely many assets. He is given a class of probability measures (which we refer to as *scenarios*), and corresponding real numbers, called *floors*. Any pay-off X that he can generate at time T is called *acceptable*, if the expectation of X under each scenario is not less than the corresponding floor. In this chapter we calculate the minimum initial capital required, so that by careful trading, following a self-financing strategy, the agent can turn the terminal value of the portfolio acceptable.

In the language of Section 2.5, we shall be computing $\tilde{\rho}(Z)$, for an arbitrary integrable random variable Z, and $\tilde{\rho}$ induced by any convex measure of risk ρ

via (2.15). The details have been worked out in Subsection 3.1.4. Hence the history and developments of the financial ideas and interpretation of results in this chapter are already described in Chapter 2. We owe the problem to a recently published paper due to Larsen, Pirvu, Shreve and Tütüncü [Larsen et al., 2004], in which the authors look at a market with a semimartingale price process and consider the same problem with finitely many scenarios. We borrow much of the mathematical structure from their paper, and extend their results to any arbitrary (even uncountable) class of scenarios. What makes our generalisation necessary is that most such natural measures would require infinitely many scenarios, and the finite-dimensional arguments used in [Larsen et al., 2004] fail to extend. We provide an example in Section 4.4 where we compute capital requirement to efficiently hedge a contingent claim when we allow controlled shortfall. This problem can easily be cast in our set-up discussed above, but only with infinitely many scenarios and floors.

Remark. We should mention here that we differ from [Larsen et al., 2004] at another important point; we only look at a *static* case, i.e., the risk is measured when we start trading at $t = 0$.

This chapter is divided as follows. Subsection 3.1.2 describes our mathematical set-up, leading to a precise statement of the problem in Subsection 3.1.3. Our solution is a reflection of the central result in [Larsen et al., 2004] for the finite case, i.e., the minimum capital required to achieve acceptability is equal to the supremum of the floors corresponding to such convex combinations of the scenarios under which the price process is a martingale. This is established through three results in Section 3.3, Propositions 3.2, 3.4, and 3.5, with increasing ease of application at the cost of generality. Several novel functional analytic and probabilistic instruments have been applied in the process. For example, Proposition 3.2, which follows from Hilbert space arguments, is a general condition but difficult to verify in practice. We use results about nearest point projections in $\mathbf{L}^p$ spaces, to obtain a more amenable one in Proposition 3.4. Proposition 3.5 exhibits how such a condition can be achieved by a proper choice of the underlying filtration. As mentioned before, an example has been worked out in Section 4.4 as an application of the theory.

3.1.1 Acknowledgements

This chapter uses mathematical theory which I was unaware of to begin with. I sincerely thank Prof. Simeon Reich, Prof. Heinz Bauschke and Prof. Leonard Gross for suggesting me proper directions and references when I needed them. I am also indebted to Prof. Hans Föllmer and Prof. Martin Schweizer for pointing out several related references.

3.1.2 Description of the market

The market we consider has one risky asset and zero risk-free interest rate. These are simplifying assumptions, not difficult to avoid. But we adhere to them for notational simplicity. The price of our risky asset is assumed to be a real-valued (although only notational changes are required, in order to handle a vector-valued semimartingale) special semimartingale $S_t,\ 0 \leq t \leq T$, adapted to a suitable filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, P)$. We assume that the filtration is right-continuous, $\mathcal{F}_0$ contains all the P null sets, and that $\mathcal{F}_T$ is the entire σ-algebra $\mathcal{F}$. The semimartingale S has the following Doob-Meyer decomposition

$$S_t = S_0 + M_t + A_t, \ \ 0 \leq t \leq T, \tag{3.1}$$

where the process M is a local martingale and A is a predictable process of finite variation. They are both assumed to be càdlàg. Without loss of generality, at time zero, the initial price S_0 is assumed to be zero.

For any special semimartingale X which can be decomposed as $X = N+V$, where N is a local martingale and V is a predictable finite variation process, one can define the $\mathcal{H}^2$ norm of X by $\|X\|^2_{\mathcal{H}} := \mathrm{E}\left([N]_T\right) + \mathrm{E}\left(|V|^2_T\right)$. Here $[N]$ is the *quadratic variation* of the local martingale N. The class of special semimartingales with a finite $\mathcal{H}^2$ norm is a Banach space (see [Protter, 2004]).

Assumption 3.1. We shall assume that the $\mathcal{H}^2$-norm of the semimartingale S in (3.1) is finite, i.e., $\|S\|_{\mathcal{H}} := \sqrt{\mathrm{E}\left([M]_T\right) + \mathrm{E}\left(|A|^2_T\right)} < \infty$.

Let Θ denote the collection of predictable processes π such that

$$\mathrm{E}\left(\int_0^T \pi_u^2 \mathrm{d}[M]_u\right) + \mathrm{E}\left(\int_0^T |\pi_u| \mathrm{d}|A|_u\right)^2 < \infty. \tag{3.2}$$

Then for any predictable process $\pi \in \Theta$ and for any $0 \leq t \leq T$, the stochastic integral of π with respect to the process S is well defined in the interval $[0,t]$ and will be denoted by $(\pi.S)_t := \int_0^t \pi_u \mathrm{d}S_u$. The process $(\pi.S)_t$ is again a special semimartingale in the interval $[0,T]$ with a finite $\mathcal{H}^2$ norm, whose square is given by (3.2). See [Protter, 2004] for the proofs.

For $x \in \mathbb{R}$ and $\pi \in \Theta$, we call

$$W_u^{x,\pi} := x + (\pi.S)_u, \quad 0 \leq u \leq T, \tag{3.3}$$

the wealth process at time u starting with initial capital x and generated by the trading strategy π. We shall make use of the following notation.

Notation 3.1. Let $\mathbf{L}^p$, $1 \leq p < \infty$, denote the space of all $\mathcal{F}$-measurable random variables which have integrable pth moment under P.

Now for any $\pi \in \Theta$, it is again a standard fact that the random variable $(\pi.S)_T$ is square integrable. Thus there is an obvious map from Θ into $\mathbf{L}^2$ which carries π to the stochastic integral $(\pi.S)_T$. We consider the range of this map

$$G := \left\{ X \in \mathbf{L}^2 \,\middle|\, X = \int_0^T \pi_u \mathrm{d}S_u, \;\; \textit{for some} \;\; \pi \in \Theta \right\}. \tag{3.4}$$

Clearly, G is a subspace of the Hilbert space $\mathbf{L}^2$. We shall denote by $\overline{G}$, the closure of G in $\mathbf{L}^2$. This closure will then be a Hilbert space in its own right.

3.1.3 Statement of the problem

Let Δ be a collection of probability measures on $(\Omega, \mathcal{F})$, which are absolutely continuous with respect to P and, let ϕ be a mapping from Δ into $\mathbb{R}$.

Assumption 3.2. Assume that

$$\Lambda \triangleq \left\{ \mathrm{d}Q/\mathrm{d}P \,\middle|\, Q \in \Delta \right\} \tag{3.5}$$

is a subset of $\mathbf{L}^2$.

Problem 3.1. Let Γ be the subset of $\mathbf{L}^2$ defined by

$$\Gamma \triangleq \left\{ X \in \mathbf{L}^2 \;\middle|\; \mathrm{E}^Q(X) \geq \phi(Q), \;\; \forall\, Q \in \Delta \right\}. \tag{3.6}$$

A real number x will be called *acceptable* if

$$(x + G) \cap \Gamma \neq \emptyset, \tag{3.7}$$

where the subspace G is defined in (4.19). We shall denote the set of acceptable initial positions by $\mathcal{A}_0$. That is to say, $x \in \mathcal{A}_0$ if there exists a $\pi \in \Theta$ such that

$$\mathrm{E}^Q\left(x + \int_0^T \pi_u \mathrm{d}S_u \right) \geq \phi(Q), \;\; \forall\, Q \in \Delta. \tag{3.8}$$

It is immediate that the set $\mathcal{A}_0$ is a half-line, unbounded from above. We shall be concerned with determining

$$\inf\{x \in \mathbb{R} |\ x \in \mathcal{A}_0\}. \tag{3.9}$$

REMARK. Another important question is whether the set $\mathcal{A}_0$ is closed or not. That is to say, whether the infimum in (3.9) is attained. As we shall see in the beginning of Section 3.3, our present set-up is deficient in answering the question. We shall get, however, a partial solution.

3.1.4 Relations with market measures of risk.

Market measures of risk has been introduced and developed (although in discrete time for simplicity) in Subsection 2.5. In this subsection we show how minimum capital requirement in our sense is equivalent to computing measures of risk in the financial market.

We start with the space $\mathbf{L}^\infty$ of real-valued, P-essentially-bounded measurable functions defined on $(\Omega, \mathcal{F}, P)$. Define the function ρ by

$$\rho(X) = \sup_{Q \in \Delta} \left(\mathrm{E}^Q[-X] + h(Q)\right), \quad X \in \mathbf{L}^\infty, \tag{3.10}$$

where Δ is as in the last section and $h : \Delta \to \mathbb{R}$. Then ρ is a convex measure of risk. Moreover, as shown in [Föllmer and Schied, 2004, page 172] (see Section 2.4, Theorem 2.2), all convex measures of risk (with some regularity) can be represented as above for some penalty function h (although Δ might not be a subset of $\mathbf{L}^2$).

Given the subspace G of (4.19), we define the market measure of risk as another measure of risk

$$\rho_G(X) \triangleq \inf_{H \in G} \rho(X - H),\ X \in \mathbf{L}^\infty.$$

This is an example of *inf-convolution* of risk measures as seen in Example 2.7 (see (2.20)).

By Assumption 3.2, we can extend the domain of ρ and ρ_G to the whole of $\mathbf{L}^2$. Fix $\chi \in \mathbf{L}^2$. We shall show that by a suitable choice of ϕ, the value of the infimum in (3.9) is equal to $\rho_G(\chi)$. To see this, define $\phi(Q) := h(Q) - \mathrm{E}^Q(\chi)$. Then, by definition (3.7), we get $\inf\{x \in \mathbb{R} \mid x \in \mathcal{A}_0\}$

$$
\begin{aligned}
&= \inf \left\{ x \in \mathbb{R} \;\middle|\; \exists\, \xi \in G,\;\; \mathrm{E}^Q(x+\xi) \geq \phi(Q),\; \forall Q \in \Delta \right\} \\
&= \inf \left\{ x \in \mathbb{R} \;\middle|\; \exists\, \xi \in G,\;\; \mathrm{E}^Q(x+\xi) \geq h(Q) - \mathrm{E}^Q(\chi),\; \forall Q \in \Delta \right\} \\
&= \inf \left\{ x \in \mathbb{R} \;\middle|\; \exists\, \xi \in G,\;\; \mathrm{E}^Q(-(\chi+x+\xi)) + h(Q) \leq 0,\; \forall Q \in \Delta \right\} \\
&= \inf \left\{ x \in \mathbb{R} \;\middle|\; \exists\, \xi \in G,\;\; \rho(\chi+x+\xi) \leq 0 \right\} \\
&= \inf \left\{ x \in \mathbb{R} \;\middle|\; \inf_{H \in G} \rho(\chi + x - H) \leq 0 \right\} && (3.11) \\
&= \inf \left\{ x \in \mathbb{R} \;\middle|\; \rho_G(\chi + x) \leq 0 \right\} = \rho_G(\chi). && (3.12)
\end{aligned}
$$

The equality in (3.11) requires proper assumption on the regularity of ρ and the last one is due to translation invariance of the convex risk measure ρ_G, see page 155, eqn.(4.5) of [Föllmer and Schied, 2004].

3.2 A general Hilbert space problem

Let $\mathcal{H}$ be a Hilbert space with an inner product denoted by $\langle .\, , .\, \rangle$ and the norm by $\|.\|$. Suppose we are given a set $\Lambda \subseteq \mathcal{H}$, a mapping $f : \Lambda \to \mathbb{R}$, and a closed subspace $\mathcal{G} \subseteq \mathcal{H}$. For any given real number x, we want to find necessary and sufficient conditions for the existence of an element $z^* \in \mathcal{G}$ such that

$$\langle z^*, y \rangle \geq f(y) - x, \;\; \forall y \in \Lambda. \tag{3.13}$$

The rest of this section is devoted to solving this problem.

Let $\widetilde{\Lambda}$ denote the convex hull of Λ. Extend the mapping f from Λ to $\widetilde{\Lambda}$ by defining a new mapping $\tilde{f} : \widetilde{\Lambda} \to \mathbb{R}$ (the *least concave majorant*), given by

$$\tilde{f}(y) := \sup \left\{ \sum_{i=1}^{n} \lambda_i f(z_i) \;\middle|\; y = \sum_{i=1}^{n} \lambda_i z_i, \quad z_i \in \Lambda,\ \lambda_i \geq 0,\ \sum_i \lambda_i = 1 \right\}. \tag{3.14}$$

Let us observe that if z^* is a solution for (3.13), then z^* also solves a more general class of inequalities. In fact, by the linearity of inner products, it follows from (3.13) that if $y \in \widetilde{\Lambda}$ can be written as as a convex combination of some $\{z_1, z_2, \ldots, z_n\} \subseteq \Lambda$, i.e. $y = \sum \lambda_i z_i$, then $\langle z^*, y \rangle = \sum \lambda_i . \langle z^*, z_i \rangle \geq \sum \lambda_i f(z_i) - x$. Thus, we can appeal to the definition of $\tilde{f}$ in (3.14) to obtain

$$\langle z^*, y \rangle \geq \tilde{f}(y) - x, \quad \forall y \in \widetilde{\Lambda}. \tag{3.15}$$

Let $\mathcal{T}(y)$ for any $y \in \mathcal{H}$ denote the unique orthogonal projection of y on $\mathcal{G}$. In particular, we have

$$\langle z, y \rangle = \langle z, \mathcal{T}(y) \rangle, \quad \forall z \in \mathcal{G}. \tag{3.16}$$

Proposition 3.1. *For any $x \in \mathbb{R}$, a necessary and sufficient condition for the existence of $z^* \in \mathcal{G}$ satisfying the inequalities in (3.13), is the existence of a constant $M \geq 0$ such that*

$$M \left\| \mathcal{T}(y) \right\| \geq \tilde{f}(y) - x, \quad \forall y \in \widetilde{\Lambda}. \tag{3.17}$$

PROOF. To see the necessity of condition (3.17), just apply the Cauchy-Schwarz inequality to (3.15) to get

$$\tilde{f}(y) - x \leq \langle z^*, y \rangle = \langle z^*, \mathcal{T}(y) \rangle \leq \|z^*\| \left\| \mathcal{T}(y) \right\|, \quad \forall y \in \widetilde{\Lambda}. \tag{3.18}$$

Setting $M := \|z^*\|$ we have established condition (3.17).

Proving the sufficiency is more subtle. We start with the assumption that (3.17) holds for some $x \in \mathbb{R}$ and some real constant $M \geq 0$. To simplify notation, let us define a new mapping $b : \widetilde{\Lambda} \to \mathbb{R}$ by

$$b(y) := \tilde{f}(y) - x, \quad y \in \widetilde{\Lambda}.$$

Condition (3.17) then reads

$$M \, \|\mathcal{T}(y)\| \geq b(y), \quad \forall \, y \in \widetilde{\Lambda}. \tag{3.19}$$

We shall establish (3.13) by showing that there exists $z^* \in \mathcal{H}$ such that

$$\langle z^*, y \rangle \geq b(y), \quad \forall \, y \in \widetilde{\Lambda}. \tag{3.20}$$

• We shall first show that for any given finite subset $\{y_1, \ldots, y_n\} \subseteq \widetilde{\Lambda}$, there is a $z^* \in \mathcal{G}$ such that $\|z^*\| \leq M$ and z^* satisfies

$$\langle z^*, y_k \rangle \geq b(y_k), \quad \forall \, 1 \leq k \leq n. \tag{3.21}$$

We shall argue this by contradiction. Suppose that no such z^* exists. Consider the set

$$\mathcal{S} := \Big\{ (\, \langle z, y_1 \rangle, \ldots, \langle z, y_n \rangle \,) \,\Big|\, z \in \mathcal{G}, \; \|z\| \leq M \Big\}$$

which is compact and convex in $\mathbb{R}^n$. Here and throughout, $\mathbb{R}^n_+$ will refer to the subset of points in $\mathbb{R}^n$ which have all co-ordinates non-negative. Let $\mathcal{S}^-$ be the set all points $(a_1, a_2, \ldots, a_n)$ which can be represented as

$$a_k = \langle z, y_k \rangle - r_k, \quad 1 \leq k \leq n,$$

for some $r_k \geq 0$ and some $z \in \mathcal{G}$ such that $\|z\| \leq M$. For notational simplicity, let us denote $b_k := b(y_k), \;\; 1 \leq k \leq n$. Since we have assumed that no solution to (3.21) exists, we have

$$(b_1, \ldots, b_n) \notin \mathcal{S}^-. \tag{3.22}$$

But, by the Separating Hyperplane Theorem, (3.22) implies that there exists a vector $\lambda = (\lambda_1, \ldots, \lambda_n) \in \mathbb{R}^n$, $\lambda \neq 0$, such that for all $a_1 \geq 0, \ldots, a_n \geq 0$ we have

$$\sum \lambda_i b_i > \sum \lambda_i \langle z, y_i \rangle - \sum \lambda_i a_i, \quad \forall z \in \mathcal{G}, \ \|z\| \leq M. \tag{3.23}$$

For any i, let a_i tend to infinity to see that $\lambda_i \geq 0$. We can thus normalize λ to have $\sum_{i=1}^n \lambda_i = 1$. Taking $a_1 = \ldots = a_n = 0$ in (3.23), we obtain

$$\sum \lambda_i b_i > \sum \lambda_i \langle z, y_i \rangle = \left\langle z, \sum \lambda_i y_i \right\rangle = \langle z, y_c \rangle \,, \tag{3.24}$$

where $y_c := \sum \lambda_i y_i \in \widetilde{\Lambda}$.

Note that the function $\tilde{f}$ in (3.14) is concave. This has been proved in [Rockafellar, 1997], page 37, Theorem 5.6, and the example following right after its proof. Thus the function $b(.) = \tilde{f}(.) - x$ is also concave. Combining with (3.19), and using the concavity of b, we get

$$M \, \|\mathcal{T}(y_c)\| \geq b(y_c) = b\left(\sum \lambda_i y_i\right) \geq \sum \lambda_i b_i > \langle z, y_c \rangle = \langle z, \mathcal{T}(y_c) \rangle, \tag{3.25}$$

for every $z \in \mathcal{G}$ with $\|z\| \leq M$. If $\|\mathcal{T}(y_c)\| = 0$, this leads to $0 \geq \sum \lambda_i b_i > 0$, which is a contradiction; whereas if $\|\mathcal{T}(y_c)\| \neq 0$, note that $z = M.\mathcal{T}(y_c) / \|\mathcal{T}(y_c)\|$ is an element of the subspace $\mathcal{G}$ with $\|z\| \leq M$ which, when plugged into inequality (3.25), gives

$$M \, \|\mathcal{T}(y_c)\| \geq \sum \lambda_i b_i > M \, \|\mathcal{T}(y_c)\| \,,$$

again a contradiction. We have thus proved (3.21).

• In general, let us define for any $y \in \widetilde{\Lambda}$, the following subset of $\mathcal{G}$:

$$\Pi_y := \left\{ z \in \mathcal{G} \;\middle|\; \|z\| \leq M, \ \langle z, y \rangle \geq b(y) \right\}.$$

Then there is a solution to (3.20) if we can show that

$$\cap_{y \in \widetilde{\Lambda}} \Pi_y \neq \emptyset. \tag{3.26}$$

Now each Π_y is a closed subset of the M-ball of $\mathcal{G}$, a set which is compact under the weak topology. This follows from the Banach-Alaoglu Theorem and the fact that a Hilbert space is its own dual (see [Rudin, 1991], pages 68, 94). Thus Π_y is a weak-compact subset of $\mathcal{G}$. Hence, by the finite intersection property, (3.26) holds if and only if for any finite collection $\{y_1, \dots, y_n\} \subseteq \widetilde{\Lambda}$, we have $\cap_{1 \le i \le n} \Pi_{y_i} \neq \emptyset$. That is to say, (3.26) holds if and only if for any finite collection $\{y_1, \dots, y_n\} \subseteq \widetilde{\Lambda}$ we can find an element $z^* \in \mathcal{H}$ such that $\|z^*\| \le M$ and $\langle z^*, y_k \rangle \ge b(y_k), \quad k = 1, \dots, n$. But this is what we have shown in (3.21). This proves the theorem. $\square$

Our previous result does not hold when the subspace $\mathcal{G}$ is not a closed subspace of the Hilbert space $\mathcal{H}$. However what we shall show now is that not much is lost if we consider $\overline{\mathcal{G}}$, the closure of $\mathcal{G}$ instead of $\mathcal{G}$ itself.

Let us denote by $\overline{\mathcal{A}}_0$ the set of all real numbers x for which the inequalities in (3.13) have a solution for some $z^* \in \overline{\mathcal{G}}$, and reserve the notation $\mathcal{A}_0$ for that subset of $\overline{\mathcal{A}}_0$ for which the solution z^* is actually an element of $\mathcal{G}$. We shall now show that when Λ is bounded in norm, $\mathcal{A}_0$ is a dense subset of $\overline{\mathcal{A}}_0$. However, since both $\mathcal{A}_0$ and $\overline{\mathcal{A}}_0$ are half-lines, this is actually equivalent to proving what we shall need most, i.e.,

$$\inf \mathcal{A}_0 = \inf \overline{\mathcal{A}}_0. \tag{3.27}$$

Lemma 3.1. *If the set Λ is bounded in norm and if $x \in \overline{\mathcal{A}}_0$, then $(x + \epsilon) \in \mathcal{A}_0$ for any positive ϵ.*

PROOF. Fix $x \in \overline{\mathcal{A}}_0$ and an $\epsilon > 0$. By the definition of $\overline{\mathcal{A}}_0$, there exists $z \in \overline{\mathcal{G}}$ such that

$$\langle z, y \rangle \ge f(y) - x, \quad \forall y \in \Lambda.$$

Now since Λ is bounded in norm and $\mathcal{G}$ is dense in $\overline{\mathcal{G}}$, there is an element $z^* \in \mathcal{G}$

such that

$$\sup_{y\in\Lambda} |\langle z^*, y\rangle - \langle z, y\rangle| \le \|z^* - z\| \,.\, \sup_{y\in\Lambda} \|y\| \le \epsilon.$$

Hence, we get $\langle z^*, y\rangle \ge \langle z, y\rangle - \epsilon \ge f(y) - (x+\epsilon), \quad \forall y \in \Lambda$. Since $z^* \in \mathcal{G}$, this shows that $(x+\epsilon) \in \mathcal{A}_0$, and proves the lemma. □

REMARK. Equation (3.27) does not hold in full generality, although we always have $\inf \mathcal{A}_0 \ge \inf \overline{\mathcal{A}}_0$, since $\mathcal{A}_0 \subseteq \overline{\mathcal{A}}_0$. We shall return to discuss this point again in the next section.

3.3 Main results

We shall now translate the results of the last subsection in order to solve (3.9). Consider the Hilbert space $\mathbf{L}^2$ and the subspace G of stochastic integrals defined in (4.19). Let $\overline{G}$ denote the closure of G in $\mathbf{L}^2$. Recall the statement of the problem in subsection 3.1.3, and as in the setting of the last section, define:

$$\mathcal{H} = \mathbf{L}^2, \quad \mathcal{G} = \overline{G}, \quad \Lambda = \left\{ \mathrm{d}Q/\mathrm{d}P \;\middle|\; Q \in \Delta \right\}. \tag{3.28}$$

That the set Λ is a subset of $\mathcal{H}$ is a consequence of Assumption 3.2. As before, $\widetilde{\Lambda}$ will denote the convex hull of Λ. Note that there is a one-to-one correspondence between the elements in Λ and the probability measures in Δ. Define the function $f : \Delta \to \mathbb{R}$ by

$$f(X) = \phi(Q), \quad \text{for } X = \mathrm{d}Q/\mathrm{d}P, \; Q \in \Lambda. \tag{3.29}$$

Define $\tilde{f}$ on $\widetilde{\Lambda}$ in the same way as in (3.14). The notation for $\|.\|$, from now on, is strictly reserved for the $\mathbf{L}^2$ norm.

Clearly with this set-up, for any $X \in \overline{G}$ and any measure Q such that $\mathrm{d}Q/\mathrm{d}P \in \widetilde{\Lambda}$, one has $\mathrm{E}^Q(X) = \langle X, \mathrm{d}Q/\mathrm{d}P\rangle$. This association makes evident the relation between solving inequalities (3.8) and (3.13). In fact, if G is a closed subspace of $\mathbf{L}^2$, solving for (3.8) is exactly the same as solving for (3.13). Problems arise when G is not closed; for then the solution obtained in (3.13) might be an element strictly in the closure of G. This problem is easy to deal with when Λ is bounded in norm, since our object of interest, $\inf \mathcal{A}_0$, remains the same whether we consider G or $\overline{G}$, as shown by Lemma 3.1 at the end of the last section.

In general, however, we cannot expect that the wealth process $\int_0^t \pi_u \mathrm{d}S_u$ which satisfies inequalities (3.8) will have finite $\mathcal{H}^2$ norms. A good analogy will be to think of situations where the optimal wealth process is a strict local martingale instead of being a true martingale. Our subspace G only allows terminal wealth from a wealth process which has finite $\mathcal{H}^2$-norm, and this is usually a strong requirement. Thus it seems necessary that we reformulate Problem 3.1 by allowing solutions which belong to $\overline{G}$ rather than G itself. We now restate Problem 3.1 in the following way:

Problem 3.2. Define the set of weakly acceptable initial positions by

$$\overline{\mathcal{A}}_0 \triangleq \left\{x \in \mathbb{R} \mid (x + \overline{G}) \cap \Gamma \neq \emptyset\right\}, \tag{3.30}$$

where the subspace G is defined in (4.19) and Γ is defined in (3.6). As before $\overline{\mathcal{A}}_0$, which is still a half-line not bounded above, is determined (up to closure) by $\inf \overline{\mathcal{A}}_0$.

The operator $\mathcal{T}$ will denote projection onto the subspace $\overline{G}$. That is, for any $X \in \mathbf{L}^2$, one has the following decomposition:

$$X = \mathcal{T}(X) + [I - \mathcal{T}](X), \tag{3.31}$$

where $[I - \mathcal{T}](X)$ is orthogonal to every element in $\overline{G}$. The following proposition is a restatement of Proposition 3.1.

Proposition 3.2. *Under Assumptions 3.1 and 3.2, a real number x is weakly acceptable (in the sense of Problem 3.2) if and only if there exists a non-negative* *real constant M such that*

$$M \|\mathcal{T}(X)\| \geq \tilde{f}(X) - x, \quad \forall\, X \in \widetilde{\Lambda}. \tag{3.32}$$

The probabilistic interpretation of $\mathcal{T}(X)$ will be clear in the next lemma.

Lemma 3.2. *For any $X \in \mathbf{L}^2$, consider the process $X_t \triangleq \mathrm{E}\left(X \,\middle|\, \mathcal{F}_t\right)$, $0 \leq t \leq T$. Then $\mathcal{T}(X) = 0$ implies $\{X_t.S_t, \mathcal{F}_t\}_{0 \leq t \leq T}$ is a martingale.*

PROOF. For any stopping time σ taking values in $[0, \tau]$, consider the process $\pi_u := 1_{\{\sigma \geq u\}}, \quad 0 \leq u \leq \tau$. Since $\mathcal{T}(X) = 0$, we have $\mathrm{E}\left(X \int_0^T \pi_u \mathrm{d}S_u\right) = \mathrm{E}\left(XS_\sigma\right) = 0$. Thus $\mathrm{E}\left(XS_\sigma\right)$ is zero for all stopping times σ. By taking conditional expectation with respect to $\mathcal{F}_\sigma$, we have $\mathrm{E}(X_\sigma S_\sigma) = 0$ for all stopping times σ. This proves the lemma. □

Lemma 3.3. *For any $Q \in \Delta$, let $\mathcal{Z}_Q = \mathrm{d}Q/\mathrm{d}P$ denote the Radon-Nikodym derivative of Q with respect to P. Then $\mathcal{T}(\mathcal{Z}_Q) = 0$ if and only if the process S is a Q-martingale on the interval $[0, T]$.*

PROOF. The *only if* part follows from the last lemma via what is commonly know as the Bayes rule. See, for example, [Karatzas and Shreve, 1991], page 193.

For the *if* part, start with a measure Q such that the process S_t is a martingale on the interval $[0, T]$. One can show by an application of the Burkholder-Davis-Gundy inequality (a proof can be found in [Larsen et al., 2004], Proposition 1) that under Assumption 3.1, for any $\pi \in \Theta$, the process $\int_0^t \pi_u \mathrm{d}S_u$ is a martingale under Q. Thus

$$\mathrm{E}\left[\frac{\mathrm{d}Q}{\mathrm{d}P} \int_0^T \pi_u \mathrm{d}S_u\right] = \mathrm{E}^Q\left(\int_0^T \pi_u \mathrm{d}S_u\right) = 0, \quad \forall \pi \in \Theta.$$

This shows that $\mathcal{Z}_Q$ is orthogonal to G and hence $\mathcal{T}(\mathcal{Z}_Q) = 0$. □

Hereafter a *martingale measure* will refer to a probability measure Q under which the process $\{S_u\}$ is a martingale in the interval $[0, T]$. Let $G^\perp$ denote the orthogonal complement of the subspace G, defined in (4.19). Then, by what we have just proved in the last lemma, the set

$$\mathcal{Z} := \widetilde{\Lambda} \cap G^{\perp} = \left\{ X \in \widetilde{\Lambda} \,\middle|\, \mathcal{T}(X) = 0 \right\} \tag{3.33}$$

is the set of probability densities corresponding to the martingale measures in the convex hull of Δ.

From now on we shall also assume the following.

Assumption 3.3. The function $\tilde{f}$ is contiuous with respect to the $\mathbf{L}^2$ norm on $\widetilde{\Lambda}$.

One can then extend $\tilde{f}$ continuously to the closure of Λ, which we shall, by an abuse of notation, continue to denote by $\widetilde{\Lambda}$. Our next theorem considers the case when $\mathcal{Z} = \emptyset$, while the other case is taken up in Proposition 3.4.

Proposition 3.3. *Suppose $\mathcal{Z} = \emptyset$. If we have $\sup_{Q \in \Delta} \phi(Q) < \infty$, or equivalently,*

$$\sup_{X \in \Lambda} f(X) < \infty, \tag{3.34}$$

then for $\inf \overline{\mathcal{A}}_0 = -\infty$, where $\overline{\mathcal{A}}_0$ is defined in Problem 3.2.

In other words, under the condition (3.34), the non-existence of martingale measures in the closed convex hull of the set of scenarios, Δ, implies that every $x \in \mathbb{R}$ is an weakly acceptable initial position.

PROOF. The set $\mathcal{T}(\widetilde{\Lambda})$, the image of $\widetilde{\Lambda}$ under the orthogonal projection mapping $\mathcal{T}$, is closed and convex. Since $\mathcal{Z} = \emptyset$, we have $0 \notin \mathcal{T}(\widetilde{\Lambda})$. Thus a basic fact from Hilbert space theory states that there is an element in $\mathcal{T}(\widetilde{\Lambda})$ which is of minimum positive norm. That is, there is an element $X^* \in \widetilde{\Lambda}$ such that $0 < \|\mathcal{T}(X^*)\| = \inf_{X \in \widetilde{\Lambda}} \|\mathcal{T}(X)\|$.

Note that (3.34) implies $\sup_{X \in \widetilde{\Lambda}} \tilde{f}(X) < \infty$. One can then define $K =$

$\max(\sup_{X\in\widetilde{\Lambda}} \tilde{f}(X) - x, 0)$, and consider $M = K/\left\|\mathfrak{T}(X^*)\right\|$ to see that

$$M\left\|\mathfrak{T}(X)\right\| \geq M\left\|\mathfrak{T}(X^*)\right\| = K \geq \tilde{f}(X) - x, \quad \forall\, X \in \widetilde{\Lambda}.$$

This shows that (3.32) is satisfied, and proves the theorem. □

For any $X \in \mathbf{L}^2$ and for any $1 \leq p \leq 2$, let us denote the $\mathbf{L}^p$ norm of X by $\left\|X\right\|_p$, i.e.,

$$\left\|X\right\|_p := \left[\mathrm{E}(|X|^p)\right]^{1/p}.$$

Since X has finite second moment and we are on a probability space, an application of Hölder's inequality shows that $\left\|X\right\|_p$ is finite for any $1 \leq p \leq 2$. We also define the $\mathbf{L}^p$-distance between a point $X \in \mathbf{L}^p$ and a non-empty subset $\Pi \subseteq \mathbf{L}^p$ by

$$d_p(X, \Pi) \triangleq \inf_{Y\in\Pi} \left\|X - Y\right\|_p. \tag{3.35}$$

Again, the distance is well-defined and finite for any $1 \leq p \leq 2$.

Proposition 3.4. *Suppose the following assumptions are satisfied:*

1. $\mathcal{Z} \neq \emptyset$.
2. *There exists a constant $L > 0$ and some $p \in (1, 2]$ such that*

$$|\tilde{f}(X) - \tilde{f}(Y)| \leq L\left\|X - Y\right\|_p \quad \forall\, X, Y \in \widetilde{\Lambda}. \tag{3.36}$$

3. *For any sequence $\{X_n\} \subseteq \widetilde{\Lambda}$ such that $\lim_{n\uparrow\infty}\left\|\mathfrak{T}(X_n)\right\| = 0$, we also have (at least through a subsequence)*

$$\lim_{n\to\infty} d_p(X_n, \mathcal{Z}) = 0. \tag{3.37}$$

Then we can conclude that

$$\inf \overline{\mathcal{A}}_0 = \sup_{Y\in\mathcal{Z}} \tilde{f}(Y). \tag{3.38}$$

Here $\overline{\mathcal{A}}_0$ is the set of weakly acceptable initial positions described in Problem 3.2, and the function $\tilde{f}$ is the least concave majorant of f defined in (3.29).

Our proof will be achieved by the following two lemmas. The first one needs the concept of *nearest point projections* in uniformly convex (or uniformly rotund) Banach spaces, e.g. the $\mathbf{L}^p$ spaces, $p \in (1, \infty)$. This can be found in [Megginson, 1998], page 427, Example 5.1.4. We can then use corollary 5.1.19 on page 435 of [Megginson, 1998], to see that given any closed, convex subset Π and any element X, both in $\mathbf{L}^p$ for some $1 < p < \infty$, there is an element $S_\Pi(X) \in \Pi$ such that

$$\|X - S_\Pi(X)\|_p = \inf_{Y \in \Pi} \|X - Y\|_p = d_p(X, \Pi). \tag{3.39}$$

Additionally, the operator S_Π is *sunny*, i.e., satisfies (see [Goebel and Reich, 1984], page 17)

$$S_\Pi\big(\alpha X + (1-\alpha) S_\Pi(X)\big) = S_\Pi(X), \;\; \forall\, \alpha \geq 0. \tag{3.40}$$

In what follows, we shall consider Π to be the closure of $\mathcal{Z}$ in $\mathbf{L}^p$. Since $\mathcal{Z}$ is $\mathbf{L}^p$ dense in Π, it follows that any real function, uniformly continuous on $\mathcal{Z}$ with respect to the $\mathbf{L}^p$ metric, can be extended uniquely on Π. By our second assumption in Proposition 3.4, the function $\tilde{f} : \widetilde{\Lambda} \to \mathbb{R}$ is uniformly continuous with respect to the $\mathbf{L}^p$ and hence can be extended to elements of Π.

The proofs of the following lemmas are done in the appendix.

Lemma 3.4. *Under the assumptions and notation of Proposition 3.4, there exists a constant $M_1 \in [0, \infty)$ such that*

$$d_p(X, \mathcal{Z}) \leq M_1 \|\mathcal{T}(X)\|, \;\; \forall\, X \in \widetilde{\Lambda}. \tag{3.41}$$

Lemma 3.5. *For a given $z \in \mathbb{R}$, suppose there exists a constant $M_2 \in [0, \infty)$ (may depend on z), such that*

$$\tilde{f}(S_\Pi(X)) - z \leq M_2 \|\mathcal{T}(X)\|, \;\; \forall\, X \in \widetilde{\Lambda}; \tag{3.42}$$

then $z \geq \sup_{X \in \mathcal{Z}} \tilde{f}(X)$. Conversely, for any $z \geq \sup_{X \in \mathcal{Z}} \tilde{f}(X)$, clearly (3.42) holds with $M_2 = 0$.

PROOF OF THEOREM 3.4. Choose $x \in \mathbb{R}$. For any $X \in \widetilde{\Lambda}$, one has the decomposition

$$\tilde{f}(X) - x = \tilde{f}(X) - \tilde{f}\left(S_{\Pi}(X)\right) + \tilde{f}\left(S_{\Pi}(X)\right) - x. \tag{3.43}$$

By Lemma 3.4, there is a $M_1 \in [0, \infty)$ such that

$$d_p(X, \mathcal{Z}) = \left\|X - S_{\Pi}(X)\right\|_p \leq M_1 \left\|\mathcal{T}(X)\right\|,$$

and thus, by assumption (2) in Proposition 3.4, we obtain

$$|\tilde{f}(X) - \tilde{f}\left(S_{\Pi}(X)\right)| \leq L \left\|X - S_{\Pi}(X)\right\|_p \leq L.M_1 \left\|\mathcal{T}(X)\right\|. \tag{3.44}$$

Plugging in the above inequality in (3.43), we see that (3.32) holds, for some $M \geq 0$, if and only if there exists a constant M_2 for which

$$\tilde{f}\left(S_{\Pi}(X)\right) - x \leq M_2 \left\|\mathcal{T}(X)\right\|, \quad \forall X \in \widetilde{\Lambda}.$$

But by Lemma 3.5, this can happen if and only if $x \geq \sup_{Y \in \mathcal{Z}} \tilde{f}(Y)$. This shows that $\inf \overline{\mathcal{A}}_0 = \sup_{Y \in \mathcal{Z}} \tilde{f}(Y)$ and proves Proposition 3.4. $\square$

Our next result displays an interesting link between the geometric and probabilistic aspects of the problem. It shows that an appropriate underlying filtration of the stock price process automatically implies (3.37).

Proposition 3.5. *Let Δ of subsection 3.1.3 be the set of all probability measures Q on $(\Omega, \mathcal{F})$, such that $Q \ll P$ and $\|\mathrm{d}Q/\mathrm{d}P\| \leq \mathcal{K}$, for some given constant $\mathcal{K} \in (0, \infty)$. As before, $\widetilde{\Lambda}$ will denote the collection of Radon-Nikodym derivatives of the measures in Δ, i.e.,*

$$\widetilde{\Lambda} \triangleq \left\{ X \in \mathbf{L}^2 \;\middle|\; X \geq 0 \text{ a.s. } P,\ \mathrm{E}(X) = 1 \text{ and } \|X\| \leq \mathcal{K} \right\}. \tag{3.45}$$

Assume that

1. *there exists an element M^* of $\mathcal{Z}$, defined in (3.33), with $\|M^*\| < \mathcal{K}$;*

2. *all martingales of the filtration $\{\mathcal{F}_t\}$ have continuous versions; and*

3. *the mapping $\tilde{f}$ satisfies (3.36).*

Then, we have $\inf \overline{\mathcal{A}}_0 = \sup_{X \in \mathcal{Z}} \tilde{f}(X)$, where $\overline{\mathcal{A}}_0$ is described in Problem 3.2, see (3.30), and the function $\tilde{f}$ is the least concave majorant of f defined in (3.29).

This theorem follows from Proposition 3.4, except we only need to show that (3.37) holds. The proof of the following lemma is in the appendix.

Lemma 3.6. *Let $\{Y_n\}$ be a sequence in $\widetilde{\Lambda}$ such that $\lim_{n\to\infty} \mathcal{T}(Y_n) = 0$. Then there exists a sequence $\{L_n\} \subseteq G^{\perp}$, with $P(L_n \geq 0) = 1$ and $\mathrm{E}(L_n) = 1$, such that:*

$$\lim_{n\to\infty} \|Y_n - L_n\|_p = 0 \tag{3.46}$$

and

$$\limsup_{n\to\infty} \|L_n\|_2 \leq \mathcal{K}. \tag{3.47}$$

PROOF OF THEOREM 3.5. Consider M^* as in assumption 1 of Proposition 3.5 and the sequence $\{L_n\}$ from Lemma 3.6. For any $\alpha \in (0,1)$, define the sequence $W_n := \alpha L_n + (1-\alpha) M^*$. Then, by the triangle inequality, we have

$$\limsup_{n\to\infty} \|W_n\| \leq \alpha \limsup_{n\to\infty} \|L_n\| + (1-\alpha) \|M^*\|.$$

Thus, from (3.47), we get that $\limsup_{n\to\infty} \|W_n\| < \mathcal{K}$. In other words, there is a large enough N such that $\|W_n\| < \mathcal{K}$ for all $n > N$. Now, by Lemma 3.6, each $P(L_n \geq 0) = 1$ and integrates to one. Also, since M^* is a probability density, $P(M^* \geq 0) = 1$ and $\mathrm{E}(M^*) = 1$. Thus we also have $P(W_n \geq 0) = 1$ and $\mathrm{E}(W_n) = 1$, and thus from (3.45), $W_n \in \widetilde{\Lambda}, \ \forall\, n > N$. But, again by Lemma 3.6, each L_n belongs to $G^{\perp}$. Since M^* also belongs to $G^{\perp}$, we conclude

$$W_n \in \widetilde{\Lambda} \cap G^{\perp} = \mathcal{Z}, \ \ \forall n > N.$$

Clearly then

$$\begin{aligned}
\limsup_{n\to\infty} d_p(Y_n, \mathcal{Z}) &\leq \limsup \|Y_n - W_n\|_p \\
&\leq \limsup \|Y_n - \alpha L_n - (1-\alpha)M^*\|_p \\
&\leq \limsup \|Y_n - L_n\|_p + (1-\alpha)\limsup \|L_n - M^*\|_p \\
&= 0 + (1-\alpha)\left(\limsup \|L_n\|_p + \|M^*\|_p\right) \\
&\leq (1-\alpha)\left(\limsup \|L_n\| + \|M^*\|\right) \leq (1-\alpha)2\mathcal{K}. \qquad (3.48)
\end{aligned}$$

The final inequality is due to (3.47) while the one right before it follows from Hölder's inequality: for any random variable Z, we have

$$\|Z\|_p \leq \|Z\|_2 = \|Z\|, \quad \forall\, 1 < p < 2. \qquad (3.49)$$

Take $\alpha \uparrow 1$ in the above inequality to conclude that $\limsup d_p(Y_n, \mathcal{Z}) = 0$ which shows (3.37) holds and the proof of Proposition 3.5 is complete. □

3.4 Examples

We solve a prototypical example of determining the sellers' price of an option in an incomplete market. Due to incompleteness of the market, a typical contingent claim will not admit a perfect hedge. Following [Föllmer and Schied, 2004], page 315, we avoid superhedging, and instead consider *efficient hedging* with controlled shortfall risk. We compute the necessary initial capital for such an efficient hedge. As expected, this capital is strictly less than the superhedging price.

Example 1. Consider a market with two stocks whose price processes S and S' are driven by a two dimensional Brownian motion till a finite terminal time T. For simplicity, the rate of interest, the mean rate of return, and the rate of dividend are kept at zero. The price process S of stock one is given by the following Black-Scholes type model:

$$\mathrm{d}S_t = S_t[\ \mu \mathrm{d}t + \sigma_1 \mathrm{d}W_1(t) + \sigma_2 \mathrm{d}W_2(t)\ ]. \tag{3.50}$$

Here the drift μ is a real constant and the volatilities σ_1, σ_2 are any two positive numbers and W_1 and W_2 are independent Brownian motions. The stochastic differential equation driving S' is left unspecified. We only assume that it is a strong solution of a differential equation involving W_1 and W_2. To generate incompleteness, we assume that trading is allowed only in stock one and not on stock two.

Now suppose we want to hedge a contingent claim C by trading in stock one. If we start with an initial investment of x and follow a trading strategy π, the wealth at the end of the trading peiod is given by

$$W_T(x, \pi) = x + \int_0^T \pi_t \mathrm{d}S_t.$$

The quantity $(C - W_T(x, \pi))^+$ is known as shortfall. In superhedging, we guarantee to have a shortfall of zero almost surely. This, however, needs a large initial amount x which sometimes investors are unable to meet. Thus it makes sense to allow shortfall in such a way that the risk is not too large.

One common way is to fix a small number α as the level of endurance and allow such strategies such that the qth. moment of the shortfall is bounded above by α. That is to say,

$$\mathrm{E}\left[(C - W_T(x, \pi))^+\right]^q \le \alpha \tag{3.51}$$

for some $q \geq 1$. Our objective is then to find the minumum real x which allows us to satisfy (3.51). Such a problem can be easily formulated as in subsection 3.1.3 by a suitable choice of convex risk measure. This has been done in detail in [Föllmer and Schied, 2004], pages 212-218, where the reader can look for the proofs.

The sample space may be any probability space Ω on which a two dimensional Brownian motion is defined. The filtration is the augmented Brownian filtration and P is the Wiener measure on this filtered probability space. We take

$$\widetilde{\Lambda} = \left\{ X \in \mathbf{L}^2 \,\middle|\, P(X \geq 0) = 1, \;\; \mathrm{E}(X) = 1 \right\} \tag{3.52}$$

and for $X \in \widetilde{\Lambda}$, define

$$\tilde{f}(X) := \mathrm{E}(XC) - (q\alpha)^{1/q} \left\| X \right\|_p . \tag{3.53}$$

Here p is given by $1/p + 1/q = 1$. We can only solve the problem for a finite q greater than 2. For such a q, it is immediate that $p \in (1, 2)$. With this definition, determining the price of the option is the same problem as stated in equation (3.8).

REMARK. We have taken $\widetilde{\Lambda}$ in (3.52) to be a subset of $\mathbf{L}^2$ which is not usual (see [Föllmer and Schied, 2004], pages 212-218). However, as long as C has more than two moments, this can be assumed without loss of generality.

First, we need to determine the elements of $\mathcal{Z}$ defined by (3.33). Since trading is allowed only on stock one, it suffices to find the probability measures in $\widetilde{\Lambda}$ under which S is a martingale in $[0, T]$. The standard tool for such problems is to use Girsanov's Theorem. Let Q be a measure equivalent to P under which S is a martingale. Without loss of generality, we can assume that

$$N_t = \mathrm{E}\left[\frac{\mathrm{d}Q}{\mathrm{d}P} \,\middle|\, \mathcal{F}_t \right] = \exp(L_t - 1/2\langle L \rangle_t), \tag{3.54}$$

for some L which is a local martingale and $\langle . \rangle$ refers to the quadratic variation of L. Then, by Girsanov's Theorem, if M_t is a martingale under the original measure P, the process $\overline{M}$, given by

$$\overline{M}_t := M_t - \langle M, L \rangle_t, \tag{3.55}$$

is a local martingale under the new measure Q. Here $\langle M, L\rangle$ refers to the mutual variation between the two processes M and L. Now the process (W_1, W_2) is a two dimensional Brownian motion. Construct a new pair of independent Brownian motions by the following rotation:

$$\widetilde{W}_1 = \frac{\sigma_1 W_1 + \sigma_2 W_2}{\sqrt{\sigma_1^2 + \sigma_2^2}}, \quad \widetilde{W}_2 = \frac{-\sigma_2 W_1 + \sigma_1 W_2}{\sqrt{\sigma_1^2 + \sigma_2^2}}$$

Clearly, $(\widetilde{W}_1, \widetilde{W}_2)$ is another two dimensional Brownian motion which generates the same filtration as (W_1, W_2). By the *Predictable Representation Property* of the Brownian filtration, one can write the local martingale L in equation (3.54) as

$$\mathrm{d}L_t = z_t \mathrm{d}\widetilde{W}_1(t) + y_t \mathrm{d}\widetilde{W}_2(t), \tag{3.56}$$

for some progressively measurable processes z and y. Thus the martingale N, in (3.54), can be written in another form

$$\mathrm{d}N_t = N_t[\ z_t \mathrm{d}\widetilde{W}_1(t) + y_t \mathrm{d}\widetilde{W}_2(t)\ ] = \mathrm{d}N_1(t) + \mathrm{d}N_2(t), \tag{3.57}$$

where N_1 and N_2 are local martingales with $\langle N_1, N_2\rangle \equiv 0$. Now, by equation (3.55), under the new measure Q, the process given by

$$\mathrm{d}\overline{W}_1(t) := \mathrm{d}\widetilde{W}_1(t) - z_t \mathrm{d}t$$

is a new Brownian motion. One can write the stochastic differential equation for S, as in equation (3.50), in terms of $\overline{W}_1$ in the following way

$$\mathrm{d}S_t = S_t[\ (\mu + \sigma^* z_t)\mathrm{d}t + \sigma^* \mathrm{d}\overline{W}_1(t)\ ]$$

where $\sigma^* = \sqrt{\sigma_1^2 + \sigma_2^2}$. Thus, if under the new measure Q, the process S is a martinagle, the only solution of z is given by

$$z_t \equiv -\mu/\sigma^*, \ \ 0 \le t \le T. \tag{3.58}$$

Since S is adapted to the filtration generated by $\widetilde{W}_1$ alone, the process y_t of (3.57) can be any progressively measurable process which makes $\int y \mathrm{d}\widetilde{W}_2$ a true martingale. This characterises the class $\mathcal{Z}$ of all the martingale measures for S.

We are now ready to solve the problems of hedging. Specifically, we take the example of the following European options with strike M, whose returns at the terminal time is

$$C = (S_T - M)^+. \tag{3.59}$$

As discussed before, we consider $\widetilde{\Lambda}$ and $\tilde{f}$ as given by (3.52) and (3.53), and try to use Proposition 3.5. We still meet some difficulties: $\widetilde{\Lambda}$ is not bounded in norm, as required by Proposition 3.5. However, we can truncate or localise the problem in the following way. For any large k let $B_k := \{X \in \mathbf{L}^2, \ \ \|X\|_2 \leq k\}$, and define

$$\widetilde{\Lambda}_k := \widetilde{\Lambda} \cap B_k, \quad \tilde{f}_k(X) := \tilde{f}(X), \ \ X \in \widetilde{\Lambda}_k. \tag{3.60}$$

Here $\widetilde{\Lambda}$ is defined in (3.52) and $\tilde{f}$ is defined in (3.53) with C as in (3.59). Intuitively, this means we are putting a heavy penalty of ∞ to measures Q which are far away from P in the sense that $\|\mathrm{d}Q/\mathrm{d}P\|_2 > k$. Now, for a large enough value of k onwards, we can assume that the set $\mathcal{Z}$ is non-empty. Also, since the random variable $(S_T - M)^+$ has all moments finite, the functional $\tilde{f}_k$ is clearly lipschitz with respect to the $\mathbf{L}^p$ norm for any $p \in (1, 2]$. We pick our favourite p to satisfy (3.36). The filtration is the augmented Brownian filtration generated by the two-dimensional Brownian motion (W_1, W_2). Thus, all martingales with respect to this filtration have continuous versions. A direct application of Proposition 3.5 would give us the following result.

RESULT. For any $y \in \mathbb{R}$, there exists a self financing trading strategy π such that

$$\mathrm{E}^Q \left(y + \int_0^T \pi_u \mathrm{d}S_u \right) \geq \tilde{f}(Q), \quad \forall Q \in \widetilde{\Lambda}_k,$$

if and only if

$$y \geq \sup_{Q \in \mathcal{Z} \cap B_k} \tilde{f}(Q). \tag{3.61}$$

Obviously, as k tends to infinity, $\widetilde{\Lambda}_k$ and $\tilde{f}_k$ tends to $\widetilde{\Lambda}$ and $\tilde{f}$ respectively. The value on the right-hand-side of (3.61) thus increases to

$$\sup_{Q \in \mathcal{Z}} \left[\mathrm{E}^Q (S_T - M)^+ - (q\alpha)^{1/q} \left\| \mathrm{d}Q/\mathrm{d}P \right\|_p \right].$$

We can define this limiting value to be the sellers' price of the option, since this is the infimum amount required to hedge the contingent claim in the sense of (3.51). Unfortunately, from our proofs it is not apparent if there is a strategy which achieves it.

3.5 Conclusion

We consider the problem of attaining acceptability by trading under convex constraints. We start with an arbitrary convex collection of scenario measures and corresponding floors, and determine the minimum capital required so that the terminal wealth can be made acceptable. Our main result states that the minimum capital is equal to the supremum of the floors over all such scenarios under which the stock price process is a martingale. We show in an example how such a result can determine the capital requirement for hedging a contingent claim with controlled shortfall.

3.6 Appendix to Chapter 3

PROOF OF LEMMA 3.4. We shall first show that for any $\epsilon > 0$, there is a $\delta > 0$ such that for any $X \in \widetilde{\Lambda}$,

$$\|\mathfrak{T}(X)\| < \delta \;\Rightarrow\; d_p(X, \mathcal{Z}) < \epsilon. \tag{3.62}$$

We shall prove this by contradiction. Fix an $\epsilon > 0$, suppose that (3.62) does not hold for any $\delta > 0$. Thus for every $\delta_n = 1/n$, one can find $X_n \in \widetilde{\Lambda}$ such that $\|\mathfrak{T}(X_n)\| < \frac{1}{n}$, but $d_p(X_n, \mathcal{Z}) \geq \epsilon$. But this is clearly impossible by (3.37).

Now take any $X \in \widetilde{\Lambda}$. Since Π is the $\mathbf{L}^p$ closure of $\mathcal{Z}$, it is clear that

$$d_p(X, \mathcal{Z}) = d_p(X, \Pi) = \|X - S_\Pi(X)\|_p . \tag{3.63}$$

From (3.40) we know that for any $\alpha \geq 0$, if we denote

$$X_\alpha := \alpha X + (1-\alpha) S_\Pi(X), \tag{3.64}$$

we have $S_\Pi(X_\alpha) = S_\Pi(X)$. Thus

$$\begin{aligned} d_p(X_\alpha, \mathcal{Z}) &= \|X_\alpha - S_\Pi(X_\alpha)\|_p = \|\alpha X + (1-\alpha)S_\Pi(X) - S_\Pi(X)\|_p \\ &= \|\alpha(X - S_\Pi(X))\|_p = \alpha \|X - S_\Pi(X)\|_p = \alpha . d_p(X, \mathcal{Z}). \end{aligned} \tag{3.65}$$

Since $S_\Pi(X)$ is an element of Π, which is the $\mathbf{L}^p$ closure of $\mathcal{Z}$, we can choose a sequence of elements $Y_n \in \mathcal{Z}$ such that $\|Y_n - S_\Pi(X)\|_p \to 0$. For an $\alpha < \delta / \|\mathfrak{T}(X)\|$ we would have

$$\|\mathfrak{T}(\alpha X + (1-\alpha)Y_n)\| = \alpha \|\mathfrak{T}(X)\| < \delta.$$

Thus, from condition (3.62) we get that

$$d_p(\alpha X + (1-\alpha)Y_n, \mathcal{Z}) < \epsilon, \; \forall \, n \in \mathbb{N}. \tag{3.66}$$

However, by the triangle inequality, we have

$$\begin{aligned} d_p(X_\alpha, \mathcal{Z}) &\leq \limsup_{n\to\infty} \Big[\|X_\alpha - \alpha X - (1-\alpha)Y_n\|_p + d_p(\alpha X + (1-\alpha)Y_n, \mathcal{Z}) \Big] \\ &\leq (1-\alpha)\limsup_n \|S_\Pi(X) - Y_n\|_p + \limsup_n d_p(\alpha X + (1-\alpha)Y_n, \mathcal{Z}) \\ &= 0 + \limsup d_p(\alpha X + (1-\alpha)Y_n, \mathcal{Z}) \leq \epsilon. \end{aligned}$$

The last inequality is due to (3.66). Thus for any $\alpha < \delta / \|\mathcal{T}(X)\|$, we have

$$d_p(X_\alpha, \mathcal{Z}) \leq \epsilon. \tag{3.67}$$

Now we prove (3.41). If $X \in \widetilde{\Lambda}$ is such that $\|\mathcal{T}(X)\| = 0$ then $X \in G^\perp$ and thus $X \in \mathcal{Z}$. Hence $X = S_\Pi(X)$ and $\|X - S_\Pi(X)\|_p = 0$ and (3.41) is obviously satisfied. If $\|\mathcal{T}(X)\| \neq 0$, choose $\alpha = \delta/(2\|\mathcal{T}(X)\|)$. Applying (3.67), we infer $d_p(X_\alpha, \mathcal{Z}) \leq \epsilon$. By taking $M_1 = 2\epsilon/\delta$, we see that (3.65) implies

$$d_p(X, \mathcal{Z}) = \frac{1}{\alpha} d_p(X_\alpha, \mathcal{Z}) \leq \frac{\epsilon}{\alpha} = M_1 \|\mathcal{T}(X)\|.$$

This proves (3.41) and hence the Lemma. □

PROOF OF LEMMA 3.5. If $X \in \mathcal{Z}$, then $\mathcal{T}(X) = 0$ and $S_\Pi(X) = X$. Thus if (3.42) holds for some $z \in \mathbb{R}$, we must have $z \geq \tilde{f}(X)$. Taking supremum over all $X \in \mathcal{Z}$, we infer $z \geq \sup_{X\in\mathcal{Z}} \tilde{f}(X)$.

Sufficiency follows, since for any $z \geq \sup_{X\in\mathcal{Z}} \tilde{f}(X) = \sup_{X\in\Pi} \tilde{f}(X)$, the left-hand side of (3.42) is non-positive, and we can take $M_2 = 0$. □

PROOF OF LEMMA 3.6. Let us remember that $\mathcal{T}$ is the projection operator onto the subspace $\overline{G}$. Thus $\mathcal{T}(Y_n) \to 0$ implies that there is a sequence $\{Z_n\} \subseteq G^\perp$, such that

$$\lim_{n\to\infty} \|Y_n - Z_n\| = 0. \tag{3.68}$$

Hence, it also follows that

$$\lim \mathrm{E}(Z_n) = \lim \mathrm{E}(Y_n) = 1. \tag{3.69}$$

Recall that $\mathrm{E}(Y_n) = 1$ for all n, simply by virtue of being a member of $\widetilde{\Lambda}$. Thus, if we define $c_n := \mathrm{E}(Z_n)$ then, by (3.69), $c_n \to 1$, and hence is non-zero for all $n > N_2$, for some $N_2 \in \mathbb{N}$. Thus, for all $n > N_2$, we can define $M_n := c_n^{-1} Z_n$ to get

$$\mathrm{E}(M_n) = 1. \tag{3.70}$$

Now, since $\sup_n \|Y_n\| \leq \mathcal{K}$ by assumption (3.45), and (3.68) holds, the sequence $\{Z_n\}$ is also uniformly bounded in the $\mathbf{L}^2$ norm. Hence, it follows that

$$\|Y_n - M_n\| \leq \|Y_n - Z_n\| + \left(1 - c_n^{-1}\right) \|Z_n\| \to 0. \tag{3.71}$$

As a corollary of the limit in (3.71), we infer that given any $\epsilon > 0$, there is a N_3 such that

$$\|M_n\| \leq \|Y_n\| + \|Y_n - M_n\| < \mathcal{K} + \epsilon, \quad \forall\, n > N_3. \tag{3.72}$$

Now $\{M_n\} \subseteq G^{\perp}$ implies $\mathcal{T}(M_n) = 0$. Thus if we define

$$M_n(t) := \mathrm{E}\left[M_n | \mathcal{F}_t\right],$$

then, by Lemma 3.2, the process

$$Y_n(t) := S(t).M_n(t) \tag{3.73}$$

is a martingale under P in the time interval $[0, T]$. Note that by assumption 2 of Proposition 3.5, we can choose a continuous version of $M_n(t)$. Since we assume $\mathcal{F}_T$ to be the entire σ-algebra, we identify

$$M_n(T) = M_n. \tag{3.74}$$

Also, by our normalisation in (3.70), we note that

$$M_n(0) = \mathrm{E}(M_n) = 1. \tag{3.75}$$

Let σ_n be the stopping time defined by

$$\sigma_n := \inf\{t \mid M_n(t) = 0\} \wedge T. \tag{3.76}$$

Claim. We shall defer the proof of the following claim:

$$\|M_n - M_n(\sigma_n)\|_p \to 0 \text{ as } n \to \infty. \tag{3.77}$$

Assuming that the above claim is true, note that, since $p > 1$, (3.77) implies

$$\lim \mathrm{E}\left(M_n(\sigma_n)\right) = \lim \mathrm{E}(M_n) = 1. \tag{3.78}$$

Thus, as before, there exists $N_4 \in \mathbb{N}$, such that for all $n > N_4$, if we define $d_n \triangleq \mathrm{E}(M_n(\sigma_n))$, then the following random variables are well-defined

$$L_n \triangleq d_n^{-1} M_n(\sigma_n), \quad \mathrm{E}(L_n) = 1. \tag{3.79}$$

Since the martingale $M_n(t)$ is continuous, by (3.75) and our choice of σ_n in (3.76), we see that

$$M_n(\sigma_n) \geq 0, \text{ a.s. } P, \tag{3.80}$$

and

$$Y_n(t \wedge \sigma_n) = M_n(t \wedge \sigma_n) S(t), \ \forall\, 0 \leq t \leq T. \tag{3.81}$$

However, by the optional sampling theorem, the process on the left-hand side of the above expression is an $\mathcal{F}_t$-martingale. Thus, the process of the right-hand side of (3.81) is also an $\mathcal{F}_t$-martingale. For every $n > N_4$, note that $d_n^{-1} M_n(t \wedge \sigma_n) = \mathrm{E}\left(L_n \mid \mathcal{F}_t\right)$, and hence

$$\{\mathrm{E}\left(L_n \,\middle|\, \mathcal{F}_t\right).S(t), \mathcal{F}_t\}_{0 \leq t \leq T} \tag{3.82}$$

is also a martingale. By (3.79) and (3.80), we can change the measure P, by defining

$$\mathrm{d}Q_n/\mathrm{d}P \triangleq L_n, \ \ \forall\, n > N_4.$$

Then from (3.82) one can use Bayes' rule in the reverse direction to conclude that Q_n is a sequence of martingale measures. Or, in other words, from Lemma 3.3, we conclude that

$$L_n \in G^{\perp}, \ \forall\, n > N_4.$$

To prove Lemma (3.6), now we only need to show that conditions (3.46) and (3.47) hold. The process $\{M_n^2(t), \mathcal{F}_t\}$ is a submartingale for every n, and hence we have

$$\|M_n(\sigma_n)\| \leq \|M_n\|\,. \tag{3.83}$$

Also from (3.78), it is immediate that $d_n \to 1$ and hence

$$\begin{aligned}\limsup_n \|L_n\| &\leq \lim d_n^{-1}.\limsup_n \|M_n(\sigma_n)\| \\ &\leq \limsup_n \|M_n\| = \limsup_n \|Y_n\| \leq \mathcal{K}.\end{aligned}$$

The only equality above is due to (3.71) and the final inequality is from (3.45). This clearly proves condition (3.47). To prove, condition (3.46), notice that, by the triangle inequality, $\lim_{n\to\infty} \|Y_n - L_n\|_p$ is bounded above by

$$\limsup \|Y_n - M_n\|_p + \limsup \|M_n - M_n(\sigma_n)\|_p + \limsup \|M_n(\sigma_n) - L_n\|_p \tag{3.84}$$

The first term is zero by (3.71). The second term is zero by (3.77). For the third term, an application of (3.49) and (3.83) will show that it is less than

$$\begin{aligned}\limsup \|M_n(\sigma_n) - L_n\| &\leq \limsup \left[\left(1 - d_n^{-1}\right)\|M_n(\sigma_n)\|\right] \\ \leq \limsup \left[\left(1 - d_n^{-1}\right)\|M_n\|\right] &= (\mathcal{K} + \epsilon)\limsup \left(1 - d_n^{-1}\right) = 0.\end{aligned}$$

The limiting bound on $\|M_n\|$ is obtained from (3.72). This proves that the left-hand side of (3.84) is zero. We have thus shown condition (3.46) holds and hence Lemma 3.6 is proved. $\square$

PROOF OF CLAIM (3.77). Finally it remains to prove (3.77). Note that by continuity of the martingale $M_n(t)$, we have $M_n(\sigma_n) = 0$ on the set $\{\sigma_n < T\}$. Also, due to (3.74), on the event $\{\sigma_n = T\}$, both the random variables M_n and $M_n(\sigma_n)$ are the same. Combining, we get

$$\mathrm{E}|M_n - M_n(\sigma_n)|^p = \mathrm{E}\left[|M_n - M_n(\sigma_n)|^p 1_{\{\sigma_n<T\}}\right] = \mathrm{E}\left[|M_n|^p 1_{\{\sigma_n<T\}}\right] \tag{3.85}$$

Fix an $\epsilon > 0$. The last term above can be expressed as:

$$\mathrm{E}\left[|M_n|^p 1_{\{\sigma_n<T\}}\right] = \mathrm{E}\left[|M_n|^p 1_{\{\sigma_n<T\}\cap\{M_n>\epsilon\}}\right] \tag{3.86}$$

$$+ \mathrm{E}\left[|M_n|^p 1_{\{\sigma_n<T\}\cap\{M_n<-\epsilon\}}\right] \tag{3.87}$$

$$+ \mathrm{E}\left[|M_n|^p 1_{\{\sigma_n<T\}\cap\{|M_n|<\epsilon\}}\right]. \tag{3.88}$$

The final term (3.88) is bounded as $\mathrm{E}\left[|M_n|^p 1_{\{\sigma_n<T\}\cap\{|M_n|<\epsilon\}}\right] \leq \epsilon^p$. The second term (3.87) can be bounded above by noting

$$\mathrm{E}\left[|M_n|^p 1_{\{\sigma_n<T\}\cap\{M_n<-\epsilon\}}\right] \leq \mathrm{E}\left[|M_n|^p 1_{\{M_n<-\epsilon\}}\right]. \tag{3.89}$$

Now, by assumption in Lemma 3.6, the sequence $\{Y_n\}$ is a sequence in $\widetilde{\Lambda}$. Hence, by (3.45), we have $P(Y_n \geq 0) = 1$. It the follows that, we have

$$P\big[|M_n| 1_{\{M_n\leq 0\}} \leq |Y_n - M_n| 1_{\{M_n\leq 0\}}\big] = 1.$$

The right-hand side of (3.89) can then be bounded above by

$$\begin{aligned}
\mathrm{E}\left[|M_n|^p 1_{\{M_n<-\epsilon\}}\right] &\leq \mathrm{E}\left[|Y_n - M_n|^p 1_{\{M_n<-\epsilon\}}\right] \\
&\leq \mathrm{E}|Y_n - M_n|^p = \left(\|Y_n - M_n\|_p\right)^p \\
&\leq \|Y_n - M_n\|^p, \text{ by (3.49)},
\end{aligned}$$

which goes to zero by (3.71). In the next paragraph, we shall show that (3.86) goes to zero. Thus, combining limits of all three terms (3.86), (3.87), and (3.88), and using (3.85), we get that $\limsup_{n\to\infty} \mathrm{E}\,|M_n - M_n(\sigma_n)|^p \le \epsilon^p$. Since the inequality above holds for all $\epsilon > 0$, we have proved (3.77).

Finally we shall show that the first term on the right-hand side of (3.88) goes to zero i.e.,

$$\lim_{n\to\infty} \mathrm{E}(|M_n|^p 1_{\{\sigma_n < T\}\cap\{M_n > \epsilon\}}) = 0. \tag{3.90}$$

For $r = 2/p$, by (3.72), we get $\sup_n \mathrm{E}(|M_n|^p)^r = \sup_n \mathrm{E}(|M_n|^2) = \sup_n(\|M_n\|)^2$ is finite. Since $r > 1$ by choice of p ($p < 2$), this shows that the random variables $\{|M_n|^p\}_{n\in\mathbb{N}}$ is uniformly integrable. Observe that the non-negative random variables

$$D_n \triangleq |M_n|^p 1_{\{\sigma_n < T\}\cap\{M_n > \epsilon\}}$$

clearly satisfy $D_n \le |M_n|^p$, for all $n \in \mathbb{N}$. Thus the collection of random variable $\{D_n\}_{n\in\mathbb{N}}$ are also uniformly integrable. Hence, to prove (3.90), it suffices to show

$$\lim_{n\to\infty} P(\{\sigma_n < T\} \cap \{M_n > \epsilon\}) = 0. \tag{3.91}$$

We shall prove (3.91) by contradiction. So, let us suppose that (3.91) does not hold, i.e., there is a $\delta > 0$ such that for a subsequence $\{n_k\} \subseteq \mathbb{N}$ we have

$$P(\{\sigma_{n_k} < T\} \cap \{M_{n_k} > \epsilon\}) > \delta, \ \ \forall\, k \in \mathbb{N}. \tag{3.92}$$

To keep notations simple, let us do away with the subsequence notation $\{n_k\}$ and assume instead

$$P(\{\sigma_n < T\} \cap \{M_n > \epsilon\}) > \delta, \ \ \forall n \in \mathbb{N}. \tag{3.93}$$

On the event $\{\sigma_n < T\}$, by the Optional Sampling Theorem, we have

$$\mathrm{E}\left(M_n | \mathcal{F}_{\sigma_n}\right) = M_n(\sigma_n) = 0, \ \ \text{a.s. } P.$$

Thus we get the following equality

$$P\left(\{\sigma_n < T\} \cap \{M_n > \epsilon\}\right) \le P\left(\{\mathrm{E}\left(M_n|\mathcal{F}_{\sigma_n}\right) = 0\} \cap \{M_n > \epsilon\}\right). \tag{3.94}$$

Define the following non-negative random variables

$$I_n := 1_{\{\mathrm{E}(M_n|\mathcal{F}_{\sigma_n})=0\}}, \quad J_n := P(M_n > \epsilon|\mathcal{F}_{\sigma_n}), \quad K_n := I_n J_n. \tag{3.95}$$

Note that by conditioning the event $\{M_n > \epsilon\}$ on $\mathcal{F}_{\sigma_n}$, we get

$$P\left(\{\mathrm{E}\left(M_n|\mathcal{F}_{\sigma_n}\right) = 0\} \cap \{M_n > \epsilon\}\right) = \mathrm{E}(I_n J_n) = \mathrm{E}(K_n),$$

forall $n \in \mathbb{N}$. Then, by (3.94) and assumption (3.93), we have

$$\mathrm{E}(K_n) > \delta, \quad \forall n \in \mathbb{N}. \tag{3.96}$$

Note that K_n is a non-negative random variable, and one can get a lower bound on the tail probability by using the following basic inequality, often known as the *second moment method*:

$$P\left(K_n \geq \frac{1}{2}\mathrm{E}(K_n)\right) \geq \frac{1}{4}\frac{\mathrm{E}(K_n)^2}{\mathrm{E}(K_n^2)}. \tag{3.97}$$

Thus, combining with (3.96), we infer

$$P\left(K_n \geq \frac{\delta}{2}\right) \geq P\left(K_n \geq \frac{1}{2}\mathrm{E}(K_n)\right) \geq \frac{1}{4}\frac{(\mathrm{E}K_n)^2}{\mathrm{E}(K_n^2)} \geq \frac{\delta}{4}. \tag{3.98}$$

Since $0 \leq K_n \leq 1$, the last inequality follows by noting that $\mathrm{E}(K_n^2) \leq \mathrm{E}(K_n)$, and hence

$$\frac{(\mathrm{E}K_n)^2}{\mathrm{E}(K_n^2)} \geq \mathrm{E}(K_n) \geq \delta.$$

Now, note that, since I_n only takes zero-one values,

$$\left\{K_n \geq \frac{\delta}{2}\right\} \;\Leftrightarrow\; \{I_n = 1\} \cap \left\{J_n \geq \frac{\delta}{2}\right\}. \tag{3.99}$$

Recall the original random variables M_n which were used to define K_n in (3.95). We denote the positive and negative parts of M_n by defining

$$M_n^+ \triangleq \max(M_n, 0) \;\text{ and }\; M_n^- \triangleq \max(-M_n, 0).$$

Then, on the set $\{I_n = 1\}$, we have $\mathrm{E}\left(M_n|\mathcal{F}_{\sigma_n}\right) = 0$, which in turn implies

$$\mathrm{E}\left(M_n^-|\mathcal{F}_{\sigma_n}\right) = \mathrm{E}\left(M_n^+|\mathcal{F}_{\sigma_n}\right), \quad \text{a.s. } P. \tag{3.100}$$

Also, on the set $\{M_n > \epsilon\}$, we obviously have $M_n = M_n^+$, and that $\{M_n^+ > \epsilon\}$. Thus on the set $\{I_n = 1\} \cap \{J_n \geq \delta/2\}$, we have

$$\begin{aligned}\mathrm{E}\left(M_n^-|\mathcal{F}_{\sigma_n}\right) = \mathrm{E}\left(M_n^+|\mathcal{F}_{\sigma_n}\right) &\geq \epsilon P(M_n^+ > \epsilon|\mathcal{F}_{\sigma_n}) \\ = \epsilon P(M_n > \epsilon|\mathcal{F}_{\sigma_n}) &= \epsilon J_n \ \geq \ \frac{\epsilon\delta}{2} \ \text{ a.s. } P.\end{aligned}$$

Combining the above inequality with (3.99) and (3.98), we get that

$$\begin{aligned}P\left(\mathrm{E}\left(M_n^-|\mathcal{F}_{\sigma_n}\right) \geq \frac{\epsilon\delta}{2}\right) &\geq P\left(\{I_n = 1\} \cap \left\{J_n \geq \frac{\delta}{2}\right\}\right) \\ &= P\left(K_n \geq \frac{\delta}{2}\right) \geq \frac{\delta}{4}, \quad \forall n \in \mathbb{N}.\end{aligned} \tag{3.101}$$

Recall the non-negative random variables Y_n as in the statement of Lemma 3.6. Note that we always have

$$M_n^- \leq (Y_n + M_n^-)1_{\{M_n^- \neq 0\}} \leq (Y_n - M_n)1_{\{M_n^- \neq 0\}} \leq |Y_n - M_n|. \tag{3.102}$$

Thus, if we let $R_n := \mathrm{E}\left(M^{n-}|\mathcal{F}_{\sigma_n}\right)$, from (3.102) we conclude $\mathrm{E}(R_n) = \mathrm{E}\left(M_n^-\right) \leq \mathrm{E}|Y_n - M_n| \leq \|Y_n - M_n\|$. And thus, by (3.71), we get $\mathrm{E}(R_n) \to 0$. But from (3.101) we get

$$P(R_n \geq \epsilon\delta/2) = P\left(\mathrm{E}\left(M_n^-|\mathcal{F}_{\sigma_n}\right) \geq \epsilon\delta/2\right) \geq \delta/4, \quad \forall n \in \mathbb{N}.$$

This clearly contradicts $\mathrm{E}(R_n) \to 0$. Thus (3.93) cannot be true and we have thus proved (3.91). This completes the proof of Claim (3.77). $\square$

Chapter 4

Computing strategies to achieve acceptability

4.1 Introduction

We start by considering a T period market model, with a single stock and a money market. To model uncertainty in stock price movements, we consider a probability space $(\Omega, \mathcal{F}, P)$ and a filtration

$$\mathcal{F}_0 \subseteq \mathcal{F}_1 \subseteq \ldots \subseteq \mathcal{F}_T \subseteq \mathcal{F}.$$

At every time point $t = 0, 1, 2, \ldots, T$, the discounted price of the stock, S_t, is assumed to be an integrable random variable measurable with respect to $\mathcal{F}_t$.

Next we consider a convex measure of risk: for every random variable X measurable with respect to $\mathcal{F}$, define

$$\rho(X) \stackrel{\triangle}{=} \sup_{1 \leq i \leq m} \left[\mathrm{E}^{Q_i}(-X) + \alpha_i \right] = \sup_{1 \leq i \leq m} \left[-\mathrm{E}(X f_i) + \alpha_i \right]. \tag{4.1}$$

Here $\{f_i \triangleq \mathrm{d}Q_i/\mathrm{d}P\}$ is a collection of Radon-Tikodym derivatives of probability measures $\{Q_i\},\ i = 1, \ldots, m$, defined on the sample space $(\Omega, \mathcal{F})$, and are absolutely continuous with respect to P. The constants $\{\alpha_i\}$ are real numbers. We restrict the above definition (4.1) to only those random variables X, for which (4.1) is well-defined.

Let us now introduce an agent who follows a self-financing portfolio by withholding ξ_t number of shares in between time periods t and $(t+1)$. Due to the non-anticipative nature of trading, we impose the natural condition that each ξ_t is an $\mathcal{F}_t$-measurable random variable. We also assume that various trading constraints dictate the existence of $\mathcal{F}_t$-measurable random variables a_t and b_t such that the agent is forced to obey

$$a_t \leq \xi_t \leq b_t, \ \ \forall\, t = 0, 1, \ldots, T-1. \tag{4.2}$$

Here $a_t < b_t$ are also assumed to be P-integrable. For any choice of initial capital w_0, and strategy $(\xi_0, \xi_1, \ldots, \xi_{T-1})$, let $V(w_0, \xi)$ denote the discounted terminal value of the portfolio, i.e.,

$$V(w_0, \xi) \triangleq w_0 + W(\xi), \quad \text{where} \tag{4.3}$$

$$W(\xi) = \sum_{t=0}^{T-1} \xi_t (S_{t+1} - S_t). \tag{4.4}$$

In this paper we investigate an algorithm to compute a near-optimal w_0 and $(\xi_0, \xi_1, \ldots, \xi_{T-1})$, such that $V(w_0, \xi)$ is acceptable, i.e.,

$$\rho(V(w_0, \xi)) \leq 0 \qquad \Leftrightarrow \qquad \rho(W(\xi)) \leq w_0.$$

The literature on convex measures of risk is almost silent about computing strategies to achieve acceptability. Although the existence of such strategies, while minimizing w_0, can be guaranteed by several results (as seen in the last chapter), to get our hands on one such strategy seems theoretically out of reach. In this chapter, therefore, we take a computational approach. Our main result, Proposition 4.1, focusses on a much smaller family of strategies, which satisfy (4.2), and is indexed by the unit simplex in $(m+1)$ dimension. The crux of the proposition is that, for a fixed w_0, if there is a strategy ξ which satisfies $\rho(V(w_0, \xi)) \leq 0$, then, without loss of generality, we can pick ξ from the smaller family.

We then devise a Monte-Carlo scheme, and directly verify non-positivity of $\rho(V(w_0, \xi))$ for strategies ξ, indexed by a fine grid in the unit simplex. We pick the one for which $\rho(V(w_0, \xi))$ is *approximately* non-positive. We give precise bounds on such approximations using combinatorial properties of the smaller family of strategies. This is done in Section 4.3. An example clarifies the process in Section 4.4.

4.1.1 Acknowledgments

I thank Prof. Peter Bank for suggesting the particular example in Section 4.4.

4.2 Main results

Denote the $(m+1)$-dimensional unit simplex by

$$\mathcal{S}_{m+1} \triangleq \left\{ y \in \mathbb{R}^{m+1} \;\middle|\; y_i \geq 0, \;\; \sum_{i=1}^{m+1} y_i = 1 \right\}. \tag{4.5}$$

Proposition 4.1. *For a fixed $w_0 \in \mathbb{R}$, let $\mathcal{L}$ be the collection of adapted processes $\xi = (\xi_0, \ldots, \xi_{T-1})$, which satisfy (4.2) and $\rho(W(\xi)) \leq w_0$. Assume that there exists*

$\zeta \in \mathcal{L}$ for which

$$\rho(W(\zeta)) < w_0, \quad \textit{and} \;\; P(\zeta_t = a_t \;\; \textit{for all} \;\; t) < 1. \tag{4.6}$$

For every $1 \leq i \leq m$, define the adapted sequence of random variables

$$v_t(f_i) \triangleq (b_t - a_t)\mathrm{E}\left[(S_{t+1} - S_t)f_i \;\middle|\; \mathcal{F}_t\right], \quad t = 0, 1, \ldots, T-1. \tag{4.7}$$

For every $\mathbf{r} \in \mathcal{S}_{m+1}$, with $r_{m+1} > 0$, consider the following weighted sum process

$$\lambda_t(\mathbf{r}) \triangleq \frac{1}{r_{m+1}} \sum_{i=1}^{m} r_i v_t(f_i), \quad t = 0, \ldots, T-1. \tag{4.8}$$

Now, let η be any mean-zero continuous probability distribution function on the real line. Then there exists a vector $\mathbf{r}^* \in \mathcal{S}_{m+1}$, with $r^*_{m+1} > 0$, such that the $\{\mathcal{F}_t\}$-adapted process

$$\xi^*_t(\omega) \triangleq (b_t - a_t)\eta(-\lambda_t(\mathbf{r}^*), \infty) + a_t, \quad t = 0, \ldots, T-1, \tag{4.9}$$

satisfies (4.2) and $\rho(W(\xi^*)) \leq w_0$.

The proof of this result will follow after we have introduced some notations. Let $[T]$ denote the set $\{0, 1, \ldots, T-1\}$. Enlarge the original sample space by considering

$$\Omega \times [T] = \Omega \times \{0, 1, \ldots, T-1\}. \tag{4.10}$$

Let $\mathcal{P}^{[T]}$ be the power set of the finite collection $\{0, 1, \ldots, T-1\}$ and let $\mathcal{F} \otimes \mathcal{P}^{[T]}$ denote the product σ-algebra between $\mathcal{F}$ and $\mathcal{P}^{[T]}$. Extract a sub σ-algebra $\widehat{\mathcal{F}}$ by defining

$$\widehat{\mathcal{F}} \triangleq \left\{ A \in \mathcal{F} \otimes \mathcal{P}^{[T]} \;\middle|\; \{\omega \; : \; (\omega, t) \in A\} \in \mathcal{F}_t, \;\; \forall \, t = 0, 1, \ldots, T-1 \right\}. \tag{4.11}$$

That $\widehat{\mathcal{F}}$ is a valid σ-algebra is straightforward to verify. Finally, let U_T denote the discrete uniform measure on $[T]$, and consider the product measure $P \otimes U_T$ on the sub σ-algebra $\widehat{\mathcal{F}}$. This gives us a probability space $\left(\Omega \times [T], \widehat{\mathcal{F}}, P \otimes U_T\right)$.

The advantages of considering the above probability space is the following lemma.

Lemma 4.1. *Let $\oplus\mathbf{L}^1(\mathcal{F}_t)$ denote the vector space of all T-vectors of the type*

$$(h_0, h_1, \ldots, h_{T-1}),$$

where each h_t is a P-integrable random variables measurable with respect to $\mathcal{F}_t$. Then there is a natural bijection T_1 between $\mathbf{L}^1\left(\Omega \times [T], \widehat{\mathcal{F}}, P \otimes U_T\right)$ and $\oplus\mathbf{L}^1(\mathcal{F}_t)$ given by

$$T_1(H)(\omega) \triangleq (H(\omega,0), H(\omega,1), \ldots, H(\omega,T-1)), \quad H \in \mathbf{L}^1\left(\Omega \times [T], \widehat{\mathcal{F}}, P \otimes U_T\right).$$

Similarly let $\oplus\mathbf{L}^\infty(\mathcal{F}_i)$ denote the vector space of all T-vectors of the type

$$(\psi_0, \psi_1, \ldots, \psi_{T-1}),$$

where each ψ_t is $\mathcal{F}_t$-measurable and P-essentially bounded. Again there is a natural bijection T_∞ between $\mathbf{L}^\infty\left(\Omega \times [T], \widehat{\mathcal{F}}, P \otimes U_T\right)$ and $\oplus\mathbf{L}^\infty(\mathcal{F}_i)$ given by

$$T_\infty(\psi)(\omega) \triangleq (\psi(\omega,0), \psi(\omega,1), \ldots, \psi(\omega,T-1)), \quad \psi \in \mathbf{L}^\infty\left(\Omega \times [T], \widehat{\mathcal{F}}, P \otimes U_T\right).$$

Proof of Lemma 4.1. We first claim that if a function L is measurable with respect to $\widehat{\mathcal{F}}$, then for each t, the random variable $L(.,t)$ is measurable with respect to $\mathcal{F}_t$. This follows from our definition of $\widehat{\mathcal{F}}$ in (4.11). One the other hand, take any vector $(l_0, l_1, \ldots, l_{T-1})$, such that each l_t is measurable with respect to $\mathcal{F}_t$, then the function

$$L(\omega,t) \triangleq l_t(\omega), \quad t = 0, \ldots, T-1,$$

is measurable with respect to $\widehat{\mathcal{F}}$. That these mappings preserve integrability and essential boundedness is straightforwrd. $\square$

For all sequence $\{\xi_t\}$ that satisfy (4.2), let us make a change of variable $\phi = \pi(\xi)$, where

$$\phi_t = \pi(\xi)_t = (\xi_t - a_t)/(b_t - a_t), \quad t = 0, \ldots, T-1, \tag{4.12}$$

then, each ϕ_t is $\mathcal{F}_t$-measurable and lies between zero and one. Denote by $\mathcal{R}$, the entire collection of ϕ processes as ξ ranges over $\mathcal{L}$, i.e.,

$$\mathcal{R} \triangleq \pi(\mathcal{L}), \tag{4.13}$$

then, by assumption (4.6), this collection is non-empty.

For any ξ, the discounted terminal value of the portfolio, W, defined in (4.4), can be expressed in terms of the $\phi = \pi(\xi)$ as

$$\begin{aligned} W(\xi) = W \circ \pi^{-1}(\phi) &= \sum_{t=0}^{T-1} (S_{t+1} - S_t)\left[(b_t - a_t)\phi_t + a_t\right] \\ &= \sum_{t=0}^{T-1} \left[(b_t - a_t)(S_{t+1} - S_t)\,\phi_t + a_t(S_{t+1} - S_t)\right] \end{aligned} \tag{4.14}$$

Thus, for any integrable f defined on $(\Omega, \mathcal{F}, P)$, one has

$$\begin{aligned} \int W(\xi) f \mathrm{d}P = \mathrm{E}(Wf) &= \sum_{t=0}^{T-1} \mathrm{E}\Big(\left[(b_t - a_t)(S_{t+1} - S_t)\,\phi_t + a_t(S_{t+1} - S_t)\right] f \Big) \\ &= \sum_{t=0}^{n-1} \mathrm{E}\left[(b_t - a_t)(S_{t+1} - S_t)\,\phi_t f\right] + \sum_{t=0}^{n-1} \mathrm{E}\left[a_t(S_{t+1} - S_t) f\right] \\ &= \sum_{t=0}^{T-1} \int_\Omega v_t(f)\phi_t \mathrm{d}P + c(f), \end{aligned} \tag{4.15}$$

where, we have named

$$v_t(f)(\omega) \triangleq (b_t + a_t)\mathrm{E}\left[(S_{t+1} - S_t)\, f \middle| \mathcal{F}_t\right](\omega) \tag{4.16a}$$

and

$$c(f) = \mathrm{E}\left[f \sum_{t=0}^{T-1} a_t(S_{t+1} - S_t) \right]. \tag{4.16b}$$

For $t \in [T]$, if we now look at ϕ and v as functions of two arguments (ω, t), i.e.,

$$\phi(\omega, t) \triangleq \phi_t(\omega), \quad v(f)(\omega, t) \triangleq v_t(f)(\omega), \quad \omega \in \Omega, \tag{4.16c}$$

then, by Lemma 4.1, both ϕ and v are $\widehat{\mathfrak{F}}$-measurable functions on $\Omega \times [T]$. Moreover, $v(f)$ is $P \otimes U_T$-integrable and $P \otimes U_T(\{0 \leq \phi \leq 1\}) = 1$, and from (4.15) we can write

$$\begin{aligned} \int W \circ \pi^{-1}(\phi) f \mathrm{d}P &= \sum_{t=0}^{T-1} \int v_t(f)\phi_t \mathrm{d}P + c(f) \\ &= T \int_{\Omega \times [T]} \phi v(f) \mathrm{d}\left(P \otimes U_T\right) + c(f). \end{aligned} \tag{4.17}$$

Proof of Proposition 4.1. Consider any mean-zero, continuous probability distribution function η on the real line. Consider the probability space $(\mathbb{R}, \mathcal{B}, \eta)$, where $\mathcal{B}$ is the Borel σ-algebra. Let Z be a random variable defined on it with distribution η.

Consider the following product space

$$\Omega \times [T] \times \mathbb{R}, \quad \widehat{\mathfrak{F}} \otimes \mathcal{B}(\mathbb{R}), \quad P \otimes U_T \otimes \eta. \tag{4.18}$$

Let us recall here that $\Omega \times [T]$, $\widehat{\mathfrak{F}}$ are defined in (4.10), (4.11), and U_T is the discrete uniform measure on the set $[T] = \{0, 1, 2, \ldots, T-1\}$.

Recall the functions f_i appearing in (4.1), and define the following functions in $\mathbf{L}^1(P \otimes U_T \otimes \eta)$:

$$g_i(\omega, t, x) \triangleq v(f_i)(\omega, t), \quad 1 \leq i \leq m, \quad g_{m+1}(\omega, t, x) \triangleq Z(x), \tag{4.19}$$

where the function v is defined in (4.16a), and (4.16c). Also define the constants

$$\gamma_i \triangleq \left(\alpha_i - w_0 - c(f_i)\right)/T, \quad i = 1, 2 \ldots, m.$$

We shall use Proposition 4.2 from the appendix with $(\widetilde{\Omega}, \mathcal{B}, \mu)$ as in (4.18), $k = m$, and $\{g_1, \ldots, g_{m+1}\}$ as defined in (4.19). As in the notation of Proposition 4.2, we define Φ to be the convex collection of all $\widehat{\mathcal{F}} \otimes \mathcal{B}(\mathbb{R})$-measurable functions ϕ such that $P \otimes U_T \otimes \eta(0 \leq \phi \leq 1) = 1$. Then, for $\psi \in \Phi$ and for $i = 1, \ldots, m$, we use Fubini's theorem to see

$$\begin{aligned}
\int_{\Omega\times[T]\times\mathbb{R}} \psi g_i \;\mathrm{d}\left(P \otimes U_T \otimes \eta\right) \geq \gamma_i \quad &\Leftrightarrow \\
\int_{\Omega\times[T]} \mathrm{E}\left(\psi \;\middle|\; \widehat{\mathcal{F}}\right) g_i \;\mathrm{d}\left(P \otimes U_T\right) \geq \left(\alpha_i - w_0 - c(f_i)\right)/T \quad &\Leftrightarrow \qquad (4.20) \\
T\int_{\Omega\times[T]} \mathrm{E}\left(\psi \;\middle|\; \widehat{\mathcal{F}}\right) v(f_i) \;\mathrm{d}\left(P \otimes U_T\right) + c(f_i) \geq \alpha_i - w_0. &
\end{aligned}$$

The first equivalence above follows, since g_i, by definition (4.19), is already $\widehat{\mathcal{F}}$ measurable.

• We shall show now that the set $\mathcal{A}$ in (4.58) is non-empty. To see this, take a process $\{\xi_t\}$ from $\mathcal{L}$, i.e., satisfies $\rho(W(\xi)) \leq w_0$. Let $\phi = \pi(\xi)$, then $\phi(\omega, t)$, as a real function on $\Omega \times [T]$, satisfies $P \otimes U_T(0 \leq \phi \leq 1) = 1$. Moreover, since $\rho(W(\xi)) \leq w_0$, expanding ρ in terms of f_i imply

$$\mathrm{E}\left(W(\xi) f_i\right) \geq \alpha_i - w_0, \quad \forall\, i = 1, 2, \ldots, m.$$

Hence, from (4.17), it follows that

$$T\int_{\Omega\times[T]} \phi v(f_i) \mathrm{d}\left(P \otimes U_T\right) + c(f_i) \geq \alpha_i - w_0, \;\; i = 1, \ldots, m.$$

Since ϕ is independent of Z, by the equivalences in (4.20), we obtain

$$\int_{\Omega\times[T]\times\mathbb{R}} \phi g_i \mathrm{d}\left(P \otimes U_T \otimes \eta\right) \geq \gamma_i, \;\; i = 1, \ldots, m, \qquad (4.21)$$

and thus the set $\mathcal{A}$ defined in (4.58) is non-empty. Thus from Proposition 4.2 we conclude the existence of an $r = (r_1, \ldots, r_{m+1}) \in \mathcal{S}_{m+1}$ such that ϕ^*, as in (4.61), satisfies $\phi^* \in \mathcal{A}$ and

$$\int \phi^* Z \mathrm{d}\left(P \otimes U_T \otimes \eta\right) \;\geq\; \int \psi Z \mathrm{d}\left(P \otimes U_T \otimes \eta\right), \quad \forall \psi \in \mathcal{A}. \tag{4.22}$$

• CLAIM. r_{m+1} is strictly positive.

To see this, suppose on the contrary $r_{m+1} = 0$. Then ϕ^* becomes independent of Z, and since Z is mean-zero by assumption, we get

$$\int \phi^* Z \mathrm{d}\left(P \otimes U_T \otimes \eta\right) = \int_{\mathbb{R}} Z \mathrm{d}\eta . \int_{\Omega \times [T]} \phi^* \mathrm{d}\left(P \otimes U_T\right) = 0. \tag{4.23}$$

Now, by (4.6), we can choose $\zeta \in \mathfrak{L}$ such that

$$\rho(W(\zeta)) < 0, \quad \text{and} \quad P(\zeta_t = a_t, \;\; \forall\, t) < 1. \tag{4.24}$$

Define the process $\sigma = \pi(\zeta)$, then, from (4.24), $P(\sigma_t = 0, \quad \forall\ t) < 1$. Now, by (4.24), and following exactly the same arguments as in (4.21), the process σ, seen as a random variable on $\Omega \times [T]$, will satisfy

$$\int_{\Omega \times [T] \times \mathbb{R}} \sigma g_i \mathrm{d}\left(P \otimes U_T \otimes \eta\right) = \int_{\Omega \times [T]} \sigma g_i \mathrm{d}\left(P \otimes U_T\right) > \gamma_i, \;\; 1 \leq i \leq m. \tag{4.25}$$

Also, from the second inequality in (4.24), we get $\left(P \otimes U_T\right)(\sigma = 0) < 1$. Together with $\left(P \otimes U_T\right)(0 \leq \sigma \leq 1) = 1$, we infer that

$$\int_{\Omega \times [T]} \sigma \mathrm{d}\left(P \otimes U_T\right) > 0. \tag{4.26}$$

For any $K > 0$, define

$$\beta_K = 1_{\{Z \geq -K\}}.$$

Note that $\lim_{K \to \infty} \eta(Z > -K) = 1$. Let essinf$(Z)$ denote the η-essential infimum of Z (can be $-\infty$). By continuity of η, and by (4.25), we can choose a K large enough, but less than $-$essinf(Z), such that $\eta(Z \geq -K) < 1$, and

$$\begin{gathered} \int_{\Omega \times [T] \times \mathbb{R}} \beta_K \sigma g_i \mathrm{d}\left(P \otimes U_T \otimes \eta\right) = \\ \eta(Z > -K) \int_{\Omega \times [T]} \sigma g_i \mathrm{d}\left(P \otimes U_T\right) > \gamma_i, \quad \forall\, 1 \leq i \leq m. \end{gathered} \tag{4.27}$$

Define a random variable $\theta : \Omega \times [T] \times \mathbb{R} \to \mathbb{R}$, by

$$\theta(\omega, t, z) = \beta_K(z)\sigma(\omega, t).$$

Clearly θ is a $\widehat{\mathcal{F}} \otimes \mathcal{B}(\mathbb{R})$-measurable and is an element of $\mathcal{A}$ by (4.27). However, note that

$$\begin{aligned} \int_{\Omega\times[T]\times\mathbb{R}} \theta Z \mathrm{d}\left(P \otimes U_T \otimes \eta\right) &= \int_{\mathbb{R}} Z 1_{\{Z \geq -K\}} \mathrm{d}\eta \int_{\Omega\times[T]} \sigma \mathrm{d}\left(P \otimes U_T\right) \\ &= \int_{-K}^{\infty} z\eta(\mathrm{d}z) \int \sigma \mathrm{d}\left(P \otimes U_T\right) > 0. \end{aligned} \tag{4.28}$$

The last inequality follows from (4.26) and the fact that for any $K < -\mathrm{essinf}(Z)$, we have

$$\int_{-K}^{\infty} z\eta(\mathrm{d}z) = \int_{0}^{\infty} z\eta(\mathrm{d}z) + \int_{-K}^{0} z\eta(\mathrm{d}z) > \int_{\mathbb{R}} z\eta(\mathrm{d}z) = 0.$$

The final equality above is a consequence of η having mean zero. Hence from (4.23), we see that

$$0 = \int \phi^* Z \mathrm{d}\left(P \otimes U_T \otimes \eta\right) < \int \theta Z \mathrm{d}\left(P \otimes U_T \otimes \eta\right),$$

which contradicts (4.22). This proves that $r_{m+1} > 0$.

From the conclusion of Proposition 4.2,we get ϕ^* to be of the form

$$\phi^* = \begin{cases} 1, \text{ if } \;\; \sum_{i=1}^{k} r_i v(f_i) + r_{m+1} Z > 0, \\ \\ 0, \text{ if } \;\; \sum_{i=1}^{k} r_i v(f_i) + r_{m+1} Z < 0. \end{cases} \tag{4.29}$$

Recall that from the definition of the function v in (4.16c), it is clear that each $v(f_i)$ is independent of Z. Now, since $r_{m+1} > 0$, we have

$$(P \otimes U_T \otimes \eta) \left(\sum_{i=1}^{k} r_i v(f_i) + r_{m+1} Z = 0 \right) =$$
$$\int \eta \left(r_{m+1} Z = -\sum_{i=1}^{k} r_i v(f_i) \;\Big|\; \widehat{\mathcal{F}} \right) \mathrm{d}\left(P \otimes U_T\right) = 0,$$

the integrand is zero being the consequence of the continuity of η and the fact that $r_{m+1} > 0$.

Thus, in (4.29), there is no loss in taking

$$\phi^* = \begin{cases} 1 & \text{if } \sum_{i=1}^k r_i v(f_i) + r_{m+1} Z > 0 \\ 0 & \text{otherwise.} \end{cases} \tag{4.30}$$

Now from Proposition 4.2 we also get that $\phi^* \in \mathcal{A}$, i.e.,

$$\int \phi^* g_i \mathrm{d}\left(P \otimes U_T \otimes \eta\right) \geq \gamma_i, \quad i = 1, 2, \ldots, m.$$

And from (4.20) this implies that

$$T \int_{\Omega \times [T]} \mathrm{E}\left(\phi^* \,\middle|\, \widehat{\mathfrak{F}}\right) v(f_i) \mathrm{d}\nu + c(f_i) \geq \alpha_i - w_0, \quad i = 1, \ldots, m, \tag{4.31}$$

where

$$\begin{aligned} \mathrm{E}\left(\phi^* \,\middle|\, \widehat{\mathfrak{F}}\right)(\omega, t) &= (P \otimes U_T \otimes \eta)\left(\sum_{i=1}^k r_i v(f_i) + r_{m+1} Z > 0 \,\middle|\, \widehat{\mathfrak{F}}\right)(\omega, t) \\ &= \eta\left(-\frac{1}{r_{m+1}} \sum_{i=1}^m r_i v(f_i)(\omega), \ \ \infty\right) = \eta\left(-\lambda_t(\omega), \infty\right), \end{aligned} \tag{4.32}$$

where the $\{\mathfrak{F}_t\}$-adapted process $\{\lambda_t\}$ is defined as in (4.8). Thus, if we let

$$\xi_t(\omega) \triangleq (b_t - a_t)\eta(-\lambda_t, \infty) + a_t, \quad t = 0, \ldots, T-1,$$

then, by (4.31) and (4.17), we conclude that $\int W(\xi) f_i \mathrm{d}P \geq \alpha_i - w_0, \ \ i = 1, \ldots, m$, or in other words, $\rho(W(\xi)) \leq w_0$. This proves the proposition. $\square$

4.3 Computations

For every $s = (s_1, \ldots, s_{m+1})$ in the unit simplex $\mathcal{S}_{m+1}$, recall from Proposition 4.1, the $\mathcal{F}_t$-adapted process

$$\lambda_t(s) \triangleq \frac{1}{s_{m+1}} \sum_{i=1}^{m} s_i v_t(f_i), \quad t = 0, 1, \ldots, T-1, \tag{4.33}$$

and the derived process

$$\xi_t(s) \triangleq (b_t - a_t)\eta(-\lambda_t(s), \infty) + a_t, \quad t = 0, \ldots, T-1. \tag{4.34}$$

For every initial w_0, Proposition 4.1 proves (under slight assumptions) the existence of an $r \in \mathcal{S}_{m+1}$ via which the process $\xi(r)$ satisfies $\rho(W(\xi(r))) \leq w_0$.

Construct a finite mesh $\mathbb{G}$ within the compact set $\mathcal{S}_{m+1}$. Suppose we can compute the value of $\rho(W(\xi(r)))$ for every $r \in \mathbb{G}$. Then for a given w_0, we can verify for each point on the grid if $\rho(W(\xi(r)))$ is less than w_0. For any $\epsilon > 0$, if the grid is fine enough, suitable smoothness assumptions will guarantee the existence of a point $r^* \in \mathbb{G}$, for which $\rho(W(\xi(r^*))) \leq w_0 + \epsilon$.

In fact, we can also optimize on w_0 by a similar procedure. For that fine mesh $\mathbb{G}$, let r^* be a grid point which attains

$$\rho(W(\xi(r^*))) = \min_{r \in \mathbb{G}} \rho(W(\xi(r))).$$

Let $w_0^* = \rho(W(\xi(r^*)))$, then, in an obvious way, the choice of $(w_0^*, \xi(r^*))$ gives a near-minimal initial capital for the problem of finding (w_0, ξ) which satisfies (4.2) and $\rho(W(\xi)) \leq w_0$.

The above procedure would work if we could theoretically compute $\rho(W(\xi(s)))$ for every $s \in \mathcal{S}_{m+1}$. This is often impossible. However, for any fixed s, we can estimate $\rho(W(\xi(s)))$ by Monte-Carlo simulations upto any desired level of accuracy.

We show in this section that it is possible to do a Monte-Carlo simulation to simultaneously approximate $\rho(W(\xi(s)))$ for *every* $s \in \mathcal{S}_{m+1}$ with a uniform error bound. This will allow our search algorithm described above to go through, up to approximations. The feasibility of our claim depends on the theory of *Uniform Law of Large Tumbers* and the related concept of Vapnik-Červonenkis dimension which is a combinatorial property of the particular structure of $\{\xi_t\}$ in (4.34). This theory is well-developed and we cherry-pick only the necessary results for our purpose. These have been stated in the appendix. Further references have also been provided for the interested reader.

Central to computing $\rho(W(\xi(s)))$, for any $s \in \mathcal{S}_{m+1}$, is to compute $\mathrm{E}(W(\xi(s)) \cdot f_i)$ for every f_i that defines ρ. Now, from equation (4.15), we can write

$$\begin{aligned}
\mathrm{E}(W(\xi(s))f_i) &= \sum_{t=0}^{T-1} \int_\Omega v_t(f_i)\eta(-\lambda_t(s),\infty)\mathrm{d}P + c(f_i) \\
&= T\int_{\Omega\times[T]} \eta(-\lambda(s),\infty)v(f_i)\mathrm{d}\left(P\otimes U_T\right) + c(f_i). \\
&= T\int_{\Omega\times[T]\times\mathbb{R}} \mathbb{I}\left\{\lambda(s)+Z>0\right\}v(f_i)\mathrm{d}\left(P\otimes U_T\otimes\eta\right) + c(f_i)
\end{aligned} \quad (4.35)$$

Here, as in the last section, Z is a random variable with law η independent of $\widehat{\mathcal{F}}$, and $\mathbb{I}\{\cdot\}$ denotes the indicator of an event.

We would now like to do a change of measure in (4.35) above with $v(f_i)$ as the 'Radon-Nikodým' derivative. This is not possibly directly, since $v(f_i)$ is not necessarily positive. However, we can work separately with $v^+(f_i) = \max(v(f_i),0)$ and $v^-(f_i) = \max(-v(f_i),0)$, which denote the positive and the negative parts respectively. Hence, one obtains

$$
\begin{aligned}
\mathrm{E}(W(\xi(s))f_i) - c(f_i) &= T \int_{\Omega\times[T]\times\mathbb{R}} \mathbb{I}\{\lambda(s)+Z>0\}\, v^+(f_i)\, \mathrm{d}\,(P\otimes U_T\otimes\eta) \\
&\quad -T \int_{\Omega\times[T]\times\mathbb{R}} \mathbb{I}\{\lambda(s)+Z>0\}\, v^-(f_i)\, \mathrm{d}\,(P\otimes U_T\otimes\eta) \\
&= a_i^+ \cdot \left(\mu_i^+ \otimes \eta\right)\{\lambda(s)+Z>0\} \\
&\quad -a_i^- \cdot \left(\mu_i^- \otimes \eta\right)\{\lambda(s)+Z>0\}.
\end{aligned} \tag{4.36}
$$

Here we have introduced several probability measures on $(\Omega\times[T],\widehat{\mathfrak{F}})$, defined by their corresponding *unnormalized* Radon-Nikodým derivatives:

$$
\mathrm{d}\mu_i^+/\mathrm{d}\,(P\otimes U_T) \propto v^+(f_i), \qquad \mathrm{d}\mu_i^-/\mathrm{d}\,(P\otimes U_T) \propto v^-(f_i), \tag{4.37a}
$$

and the corresponding constants

$$
a_i^+ \triangleq \sum_{t=0}^{T-1} \mathrm{E}[v_t^+(f_i)], \quad a_i^- \triangleq \sum_{t=0}^{T-1} \mathrm{E}[v_t^-(f_i)], \quad i=1,2,\ldots,m. \tag{4.37b}
$$

If any of the constants in (4.37b) is zero, the corresponding measure becomes the zero measure and can be dropped from our analysis. For efficiency in computation we would like to keep track of

$$
\aleph \triangleq \sum_{i=1}^{m} \left(1_{\{a_i^+>0\}} + 1_{\{a_i^->0\}}\right). \tag{4.38}
$$

Assumption 4.1. Throughout the rest of this section, we shall assume that

1. Ω is a subset of a Euclidean space and one can perfectly generate samples from the joint distribution of $(S_0, S_1, \ldots, S_T)$,

2. the random variables $v_t(f_i)$ (thus also λ_t) can be perfectly evaluated as functions of $(S_0,\ldots,S_T)$, and

3. the constants $c(f_i)$, a_i^+ and a_i^- can be evaluated for every $1 \leq i \leq m$.

Now, by (4.36), evaluating $\mathrm{E}(W(\xi(s))f_i)$ boils down to evaluating the following two probabilities

$$\left(\mu_i^+ \otimes \eta\right)\{\lambda(s) + Z > 0\}, \quad \text{and} \quad \left(\mu_i^- \otimes \eta\right)\{\lambda(s) + Z > 0\}. \tag{4.39}$$

It will be difficult to compute the quantities above for every $s \in \mathcal{S}_{m+1}$. Instead, we use the Vapnik-Červonenkis theory, described in Subsection 4.5.2 in the Appendix, to set up a Monte-Carlo scheme to estimate them for all s with uniform precision. The key to this is to observe the trivial equality

$$\{\lambda(s) + Z > 0\} = \left\{\sum_{j=1}^{m} s_j v(f_j) + s_{m+1} Z > 0\right\}. \tag{4.40}$$

We now apply Dudley's Theorem, Theorem 4.3 in the Appendix, with $X = \Omega \times [T] \times \mathbb{R}$ and the vector space G to be linear space spanned by Z and $v(f_j)$, $j = 1, 2, \ldots, m$. Thus we infer that the collection of sets

$$\left\{\left\{\tilde{\omega} \in \Omega \times [T] \times \mathbb{R} \ : \ \sum_{j=1}^{m} r_j v(f_j)(\tilde{\omega}) + r_{m+1} Z(\tilde{\omega}) > 0\right\}, \ r \in \mathbb{R}^{m+1}\right\}, \tag{4.41}$$

has a VC dimension not more than $(m+1)$. From (4.40), the collection of sets

$$\{\{\lambda(s) + Z > 0\}, \quad s \in \mathcal{S}_{m+1}\}$$

is contained in (4.41), and hence also has a VC-dimension not more than $(m+1)$. It is hence possible to estimate the probabilities in (4.39), uniformly for all $s \in \mathcal{S}_{m+1}$, by drawing independent samples from distributions $\mu_i^+ \otimes \eta$ and $\mu_i^- \otimes \eta$.

Our aim now would be to apply Theorem 4.6. We first have to choose two positive parameters, ϵ and δ, determining the precision of our estimates. Now, for every $i = 1, 2, \ldots, m$, choose κ_i^+ such that

$$4(\kappa_i^+)^{2(m+1)} \exp\left(-2\kappa_i^+ \left(\frac{\epsilon}{a_i^+}\right)^2 + 4\left(\frac{\epsilon}{a_i^+}\right) + 4\left(\frac{\epsilon}{a_i^+}\right)^2\right) \le \delta. \tag{4.42}$$

Generate κ_i^+ many iid samples $\{(\omega_j, t_j, z_j) \in \Omega \times [T] \times \mathbb{R},\ j = 1, 2, \ldots, \kappa_i^+\}$, from the joint distribution $\mu_i^+ \otimes \eta$.

Remark. It is fairly standard to generate samples from measures μ_i^+, defined through their unnormalized densities given in (4.37a). We can either directly identify the distribution, as we do in the next section. Or, under the assumption that one can generate perfect samples from the underlying distribution $(P \otimes U_T)$, one can use any of the standard Markov Chain algorithms, from the simple rejection sampling, to the general Metropolis-Hastings algorithm to generate samples from μ_i^+. Several books, e.g. [Gelman et al., 2003, Chap. 11], describe the details of all these algorithms.

Let $\mathcal{E}_i^+(\cdot)$ denote the empirical estimates of probabilities by the sample frequency. For example, for any $s \in \mathcal{S}_{m+1}$, we have

$$\mathcal{E}_i^+\{\lambda_s + Z > 0\} = \frac{1}{\kappa_i^+} \sum_{j=1}^{\kappa_i^+} \mathbb{I}\{\lambda_{t_j}(s)(\omega_j) + z_j > 0\}. \tag{4.43}$$

We can now apply (4.74) from Theorem 4.6 to claim that under the joint distribution of all the κ_i^+ many samples drawn

$$\text{Prob}\left\{\sup_{s \in \mathcal{S}_{m+1}} a_i^+ |\ \mathcal{E}_i^+\{\lambda_s + Z > 0\} - (\mu_i^+ \otimes \eta)\{\lambda_s + Z > 0\}\ | > \epsilon\right\} \le \delta,\ \forall i. \tag{4.44}$$

Exactly in the same way, one can replace the μ_i^+ by μ_i^- above, compute κ_i^- by

$$4(\kappa_i^-)^{2(m+1)} \exp\left(-2\kappa_i^- \left(\frac{\epsilon}{a_i^-}\right)^2 + 4\left(\frac{\epsilon}{a_i^-}\right) + 4\left(\frac{\epsilon}{a_i^-}\right)^2\right) \le \delta, \tag{4.45}$$

and obtain estimates $\mathcal{E}_i^-$, analogous to (4.43), which satisfies

$$\text{Prob}\left\{\sup_{s \in \mathcal{S}_{m+1}} a_i^- |\ \mathcal{E}_i^-\{\lambda_s + Z > 0\} - (\mu_i^- \otimes \eta)\{\lambda_s + Z > 0\}\ | > \epsilon\right\} \le \delta,\ \forall i. \tag{4.46}$$

From (4.44) and (4.46), it follows, by using (4.36), that one can estimate the quantity $\mathrm{E}(-W(\xi(s))f_i) + \alpha_i$ by

$$\mathcal{D}_i(s) \triangleq -a_i^+ \mathcal{E}_i^+\{\lambda(s) + Z > 0\} + a_i^- \mathcal{E}_i^-\{\lambda(s) + Z > 0\} - c(f_i) + \alpha_i. \tag{4.47}$$

Since $\rho(W(\xi(s))) = \sup_{1\le i\le m}\{\mathrm{E}(-W(\xi(s))f_i) + \alpha_i\}$, it follows that a good estimate of this quantity would be $\sup_i \mathcal{D}_i(s)$. We can sum-up this approximation by a simple union bound using (4.44) and (4.46) as follows.

Under the joint distribution of all the $\{\kappa_i^+, \kappa_i^-\}_{1\le i\le m}$ samples drawn from the distributions $\{\mu_i^+ \otimes \eta,\ \mu_i^- \otimes \eta\}_{1\le i\le m}$, one has

$$\mathrm{Prob}\left\{\sup_{s\in\mathcal{S}_{m+1}} |\sup_i \mathcal{D}_i(s) - \rho(W(\xi(s)))\ | \ge \epsilon\right\} \ge 1 - \aleph\delta.$$

Here, the empirical estimates $\mathcal{D}_i$ are given in (4.47), and the number $\aleph$ ($\le 2m$) is described in (4.38). We use the number $\aleph$ and not the crude bound $2m$ to bring more efficiency in our estimate.

Now that we have estimated $\rho(W(\xi(s)))$ for every $s \in \mathcal{S}_{m+1}$ with uniform precision, we can carry out the grid searching procedure described at the beginning of this section to get a near-optimal pair (w_0, ξ) which satisfies (4.2) and $\rho(W(\xi)) \le w_0$. The next section displays the entire method through an explicit example.

4.4 Examples

The previous theory is now applied to an explicit example where stock prices follow geometric Brownian motion, but observed only at finitely many time points.

We consider $T = 3$ and $\Omega = \mathbb{R}^T$, the σ-algebra $\mathcal{F}_t$ being generated by the first t co-ordinates of $\omega \in \Omega$. We take $\mathcal{F}_0$ to be the trivial σ-algebra $\{\emptyset, \Omega\}$. Take P

to be the product probability measure of T many independent Normal distributions with mean zero and variance one. In other words, we consider random variables $(Z_1, Z_2, \ldots, Z_T)$ such that each Z_i is independent and identically distributed as $N(0,1)$. The discounted stock price movement, under P, is described by

$$S_0 = 4, \quad S_{t+1} = S_t \exp\left[-\frac{1}{2} + Z_{t+1}\right], \; t = 0, 1, \ldots, T-1. \tag{4.48}$$

In other words, we have

$$S_t = S_0 \exp\left[\sum_{i=1}^{t} Z_k - \frac{t}{2}\right], \quad t = 0, 1, \ldots, T-1. \tag{4.49}$$

However, the investor is not entirely certain of his modeling assumptions, and so considers other scenarios Q_1 and Q_2, where Q_1 and Q_2 are two probability measures defined on $(\Omega, \mathcal{F}_T)$ by

$$\begin{aligned} &\text{under } Q_1, \qquad Z_1, \ldots, Z_T \overset{iid}{\sim} N(1,1), \\ &\text{under } Q_2, \qquad Z_1, \ldots, Z_T \overset{iid}{\sim} N(-1,1). \end{aligned}$$

For convenience we also introduce $Q_3 = P$.

Remark. Note, from (4.48), the effect of changing measure on the stock price movements. For Q_1, the geometric Brownian motion gets a positive drift, for Q_2 it gets a negative drift, while Q_3 is the same as P, where stock prices are a martingale.

Assume that various constraints dictate that his trading strategy is bounded between zero and one throughout, i.e., in the notation of (4.2), we have

$$a_t \equiv 0, \qquad b_t \equiv 1, \quad \text{for all} \quad 0 \le t \le T-1.$$

Now, the investor sets to do the following: if the conditions are favorable, and the stock prices tend to go up under Q_1, he wants a large lower bound e^4 for his terminal wealth. On the other hand, if the stock prices tend to go down, under Q_2, he sets a lower bound for his losses, by setting that his final wealth should be more than e^{-1}. He has at least \$1 to invest, and would like to know an optimal initial capital, and a trading strategy to achieve his goals.

This requires us to define a measure of risk ρ: if X is measurable with respect to $\mathcal{F}_T$, then

$$\rho(X) \triangleq \max_{i=1,2,3} [\mathrm{E}^{Q_i}(-X) + \alpha_i], \qquad m = 3,$$

where

$$\alpha_1 = e^4, \quad \alpha_2 = e^{-1}, \quad \alpha_3 = 1.$$

Then, we would like to compute a near-optimal pair (w_0, ξ) of initial capital w_0 and $0 \leq \xi_t \leq 1$, for all $0 \leq t \leq T-1$, such that

$$\rho(w_0 + W(\xi)) \leq 0 \quad \Leftrightarrow \quad w_0 + \mathrm{E}^{Q_i}[W(\xi)] \ \geq \alpha_i, \quad i = 1, 2, 3.$$

The first step will be to compute the functions f_1, f_2, and f_3. They are straight forward since

$$\begin{aligned} f_1(z_1, \ldots, z_k) &= \mathrm{d}Q_1/\mathrm{d}P = \exp\left[\sum_{k=1}^{T} z_k - T/2\right] \\ f_2(z_1, \ldots, z_k) &= \mathrm{d}Q_2/\mathrm{d}P = \exp\left[-\sum_{k=1}^{T} z_k - T/2\right] \\ f_3(z_1, \ldots, z_k) &= \mathrm{d}Q_3/\mathrm{d}P \equiv 1. \end{aligned} \tag{4.50}$$

We can now compute the functions $v_t(f_i)$. These are given by

$$\begin{aligned} v_t(f_1) &= \mathrm{E}\left[f_1(S_{t+1} - S_t) \mid \mathcal{F}_t\right] \\ &= S_t \mathrm{E}\left[f_1\left(\exp(Z_{t+1} - 1/2) - 1\right) \mid \mathcal{F}_t\right], \quad \text{from (4.48)}, \\ &= S_t \exp\left(\sum_{k=1}^{t} Z_k\right) \mathrm{E}\left(\exp\left\{\sum_{k=t+1}^{T} Z_k - T/2\right\}\left[\exp(Z_{t+1} - 1/2) - 1\right]\right), \end{aligned} \tag{4.51}$$

where the last equality is due to (4.50) and the independence of $\{Z_i\}$. Recall that if Z follows $N(0,1)$, then

$$\mathrm{E}\left[\exp(\sigma Z)\right] = \exp(\sigma^2/2), \qquad \sigma \in \mathbb{R}.$$

Thus, for $\mathbf{z} = (z_1, z_2, \ldots, z_T) \in \Omega$, a straightforward computation leads to

$$\begin{aligned} v_t(f_1)(\mathbf{z}) &= S_t \exp\left[\sum_{k=1}^{t} z_k\right] \left\{\exp\left(1 - \frac{t}{2}\right) - \exp\left(-\frac{t}{2}\right)\right\} \\ &= 4(\mathrm{e}-1)\exp\left\{2\sum_{1}^{t} z_k - t\right\}, \quad \text{by (4.49).} \end{aligned} \tag{4.52}$$

In particular, we have

$$\mathrm{E}(v_t(f_1)) = 4(\mathrm{e}-1)\mathrm{E}\left[\exp\left(2\sum_{k=1}^{t} Z_k - t\right)\right] = 6.87\mathrm{e}^t.$$

Similarly, we compute

$$\begin{aligned} v_t(f_2) &= \mathrm{E}\left[f_2(S_{t+1} - S_t) \mid \mathcal{F}_t\right] \\ &= S_t\mathrm{E}\left[f_2\left(\exp(Z_{t+1} - 1/2) - 1\right) \mid \mathcal{F}_t\right], \quad \text{from (4.48),} \\ &= S_t \exp\left(-\sum_{k=1}^{t} Z_k\right) \mathrm{E}\left(\exp\left\{-\sum_{k=t+1}^{T} Z_k - T/2\right\}\left[\exp(Z_{t+1} - 1/2) - 1\right]\right) \\ &= -S_0\exp(-t)\frac{\mathrm{e}-1}{\mathrm{e}} = -4(\mathrm{e}-1)\exp(-t-1). \end{aligned} \tag{4.53}$$

And obviously, since S_t is a martingale under Q_3, we have

$$v_t(f_3) = \mathrm{E}\left[S_{t+1} - S_t \mid \mathcal{F}_t\right] = 0.$$

Hence, for $s = (s_1, s_2, s_3, s_4) \in \mathcal{S}_4$, the random variable $\lambda_t(s)$ is given by

$$\begin{aligned} \lambda_t(s) &= \frac{4\mathrm{e}^{-t}(\mathrm{e}-1)}{s_4}\left[s_1\exp\left\{2\sum_{1}^{t} z_k\right\} - s_2\exp(-1)\right] \\ &= \frac{4\mathrm{e}^{-t}(\mathrm{e}-1)}{s_4}\left[s_1\mathrm{e}^t\left(\frac{S_t}{S_0}\right)^2 - s_2\exp(-1)\right] \\ &= 4(\mathrm{e}-1)\left[\frac{s_1}{s_4}\left(\frac{S_t}{S_0}\right)^2 - \frac{s_2}{s_4}\mathrm{e}^t\right]. \end{aligned}$$

Thus, for $1 \le t \le 2$ and $\mathbf{z} = (z_1, z_2, z_3) \in \Omega$, we have the following table:

$$
\begin{array}{llll}
v^+(f_1)(t,\mathbf{z}) = v_t(f_1)(\mathbf{z}), & a_1^+ = 76.34, & v^-(f_1)(t,\mathbf{z}) = 0, & a_1^+ = 0, \\
v^+(f_2)(t,\mathbf{z}) = 0, & a_1^+ = 0, & v^-(f_2)(t,\mathbf{z}) = 2.53\mathrm{e}^{-t}, & a_2^- = 3.80, \\
v^+(f_3)(t,\mathbf{z}) = 0, & a_3^+ = 0, & v^+(f_3)(t,\mathbf{z}) = 0, & a_3^- = 0.
\end{array}
$$

From above and (4.38), we also have

$$\aleph = 2.$$

Clearly, we need to consider only two changes of measures, the one given by $v^+(f_1)$ and the other by $v^-(f_2)$. The rest are all zero measures. Finally, since $a_t \equiv 0$, from (4.16b), we get

$$c(f_i) = 0, \qquad i = 1, 2, 3.$$

We take the precision parameters to be

$$\epsilon = .5, \qquad \delta = .05.$$

From (4.42) and (4.45), we determine a sufficient number of samples for desired accuracy would be

$$\kappa_1^+ = 1,400,000, \qquad \kappa_2^- = 10500.$$

Let us now analyze the probability measures μ_1^+ and μ_2^- on $\mathbb{R}^3 \times \{0, 1, 2\}$. If $\mathbf{z} \in \mathbb{R}^3$, and $0 \le t \le 2$, then from (4.37a) and (4.52) we get

$$
\begin{aligned}
\mathrm{d}\mu_1^+(\mathbf{z}, t) &\propto v^+(f_1)(\mathbf{z}, t) \cdot \mathrm{d}\left(P \otimes U_T\right)(\mathbf{z}, t) \\
&\propto \exp\left\{2\sum_1^t z_k - t\right\} \cdot \left(\frac{1}{\sqrt{2\pi}}\right)^3 \exp\left\{-\frac{1}{2}\sum_{k=1}^3 z_k^2\right\} \\
&\propto \mathrm{e}^{-t}\left(\frac{1}{\sqrt{2\pi}}\right)^3 \exp\left\{-\frac{1}{2}\sum_{k=1}^t (z_k-2)^2 - \frac{1}{2}\sum_{k=t+1}^3 z_k^2\right\}.
\end{aligned}
\tag{4.54}
$$

Thus generating a sample from μ_1^+ is the same as picking a $t \in (0, 1, 2)$ randomly with probability proportional to $\exp(-t)$. Then, conditionally on t, we generate t independent samples $Z_1, \ldots, Z_t$ from $N(2, 1)$, and $3 - t$ samples from $N(0, 1)$.

Simulating from μ_2^- is even simpler, since, from (4.53), we get that

$$\begin{aligned} \mathrm{d}\mu_2^-(\mathbf{z}, t) &\propto v^-(f_2)(\mathbf{z}, t) \cdot \mathrm{d}\left(P \otimes U_T\right)(\mathbf{z}, t) \\ &\propto \mathrm{e}^{-t} \cdot \left(\frac{1}{\sqrt{2\pi}}\right)^3 \exp\left\{-\frac{1}{2}\sum_{k=1}^{3} z_k^2\right\}. \end{aligned} \tag{4.55}$$

Here, we pick t from $\{0, 1, 2\}$ with probability proportional to $\exp(-t)$, and generate $(Z_1, \ldots, Z_T)$ as independent and identically distributed samples from $N(0, 1)$.

Finally, we take η to be $N(0, 1)$.

RESULT OF SIMULATIONS. An estimate of the minimum capital is $w_0 = 2.52$. And, an estimate of the trading strategy for this capital is $\xi_t^* = \Phi\left(-\lambda_t\right)$, where Φ is the standard normal cumulative distribution function, and λ_t is the process given by

$$\lambda_t = 4(\mathrm{e} - 1)\left[.38\left(\frac{S_t}{S_0}\right)^2 - .99\mathrm{e}^t\right].$$

4.5 Appendix to Chapter 4

4.5.1 A generalised Neyman-Pearson Lemma

Let $(\widetilde{\Omega}, \mathcal{B}, \mu)$ be a probability space. Let $\mathbf{L}^1(\mu)$ denote the space of all μ-integrable functions on $\widetilde{\Omega}$. Consider a subset $\{g_1, g_2, \ldots, g_k, g_{k+1}\}$ of $\mathbf{L}^1(\mu)$ for some integers k. Also consider real constants $\{\gamma_1, \gamma_2, \ldots, \gamma_k\}$.

Define Φ as the set of $\mathcal{B}$-measurable functions ϕ that satisfy $\mu\left(0 \leq \phi \leq 1\right) =$

1. Define $F : \Phi \longrightarrow \mathbb{R}^{k+1}$ by

$$F(\phi) \triangleq \left(\int \phi g_1 \mathrm{d}\mu, \int \phi g_2 \mathrm{d}\mu, \ldots, \int \phi g_{k+1} \mathrm{d}\mu \right), \quad \phi \in \Phi, \tag{4.56}$$

and let

$$T \triangleq -F(\Phi) = \{-F(\phi), \quad \phi \in \Phi\}. \tag{4.57}$$

The set T is obviously a closed, bounded (hence compact), convex subset of the $(k+1)$-dimensional Euclidean space.

Proposition 4.2. *Suppose the set*

$$\mathcal{A} \triangleq \left\{ \phi \in \Phi \;\middle|\; \int \phi g_i \mathrm{d}\mu \;\geq\; \gamma_i, \quad \forall\, i = 1, \ldots, k \right\} \tag{4.58}$$

is non-empty. Consider the function $\mathcal{T} : \Phi \to \mathbb{R}$ given by

$$\mathcal{T}(\phi) = \int \phi g_{k+1} \mathrm{d}\mu, \tag{4.59}$$

and denote the $(k+1)$-dimensional simplex by

$$\mathcal{S}_{k+1} \triangleq \left\{ y \in \mathbb{R}^{k+1} \;\middle|\; y_i \geq 0, \quad \sum_{i=1}^{k+1} y_i = 1 \right\}. \tag{4.60}$$

Then there exists a vector $r \in \mathcal{S}_{k+1}$ such that the function $\mathcal{T}$ is maximized over $\mathcal{A}$ by a function ϕ^* satisfying

$$\phi^* = \begin{cases} 1, \text{ if } & \sum_{i=1}^{k+1} r_i g_i(\omega) > 0 \\ \\ 0, \text{ if } & \sum_{i=1}^{k+1} r_i g_i(\omega) < 0. \end{cases} \tag{4.61}$$

PROOF OF PROPOSITION 4.2. Define the continuous function δ on T by

$$\delta(x) \triangleq x_{k+1},$$

which makes the following diagram commute:

$$\begin{tikzcd} \Phi \arrow[r, "F"] \arrow[dr, "\mathcal{T}"'] & T \arrow[d, "\delta"] \\ & \mathbb{R} \end{tikzcd}$$

Now, δ being a continuous function, achieves its minimum on the compact set $F(\mathcal{A})$ which is a closed subset of T. This proves that the function $\mathfrak{T}$ has a maximum in $\mathcal{A}$. It remains to prove that the maximizer can be taken of the form (4.61).

We shall follow closely to Ferguson [Ferguson, 1967], providing references wherever needed. First we need a few definitions: a set S in $\mathbb{R}^{k+1}$ will be called *bounded from below*, if there exists a finite number η, such that for every $y \in S$

$$y_j > \eta, \text{ for } j = 1, 2, \ldots, (k+1). \tag{4.62}$$

Let x be a point in $\mathbb{R}^{k+1}$. The *lower quantant* of x, denoted by Q_x, is defined as the set

$$Q_x = \{y \in \mathbb{R}^{k+1} : y_j \leq x_j \text{ for } i = 1, 2, \ldots, (m+n)\}. \tag{4.63}$$

A point x is said to be a *lower boundary point* of a convex set $S \in \mathbb{R}^{m+n}$ if

$$Q_x \cap \bar{S} = \{x\}. \tag{4.64}$$

Here $\bar{S}$ is the closure of the set S. The set of all lower boundary points of a convex set S is denoted by $\lambda(S)$. We now state a lemma whose proof can be found in Ferguson [Ferguson, 1967] p. 69.

Lemma 4.2. *If a non-empty convex set S is bounded from below, then $\lambda(S)$ is non-empty.*

Finally we shall need one of the basic properties of euclidean spaces:

Lemma 4.3. *(The Separating Hyperplane Theorem.) Let S_1 and S_2 be disjoint convex subsets of $\mathbb{R}^{m+n}$. Then there exists a vector $p \neq 0$ such that $\langle p, y\rangle \leq \langle p, x\rangle$ for all $x \in S_1$ and all $y \in S_2$. Here $\langle a, b\rangle$ refers to the inner product between vectors a and b.*

For a proof see, for instance, Ferguson [Ferguson, 1967] p. 73.

Let d^* denote the maximum that δ achieves on the set $F(\mathcal{A})$. Define a vector ν by declaring

$$\nu_i := -\gamma_i, \quad i = 1, 2, \ldots, k, \quad \text{and} \quad \nu_{k+1} := -d^*. \tag{4.65}$$

Consider Q_ν as defined in (4.63). Then the set

$$T' \triangleq T \cap Q_\nu \tag{4.66}$$

is non-empty, convex, and bounded from below. Hence by Lemma 4.2, the set $\lambda(T')$ of lower boundary points of T' is non-empty. Fix $x_0 \in \lambda(T')$. Note that, by (4.66), we get

$$x_0 \in T' \subseteq Q_\nu. \tag{4.67}$$

Also by definition (4.64), we have

$$Q_{x_0} \cap T' = \{x_0\}.$$

It follows that

$$Q_{x_0} \cap T = \{x_0\}, \tag{4.68}$$

since

$$Q_{x_0} \cap T' = Q_{x_0} \cap Q_\nu \cap T = Q_{x_0} \cap T.$$

The first equality above follows from (4.66), while the second follows from the fact that $Q_{x_0} \subseteq Q_\nu$. It is immediate from (4.68) that if we let C denote the interior of Q_{x_0}, then the sets T and C are disjoint convex sets. By Lemma 4.3, there exists a vector $p \neq 0$ such that

$$\langle\, p\, ,\; y\, \rangle \leq \langle\, p\, ,\; x_0\, \rangle \leq \langle\, p\, ,\; x\, \rangle, \quad \forall\, y \in C,\ \forall\, x \in T. \tag{4.69}$$

Take any y in C and choose a coordinate i. Let $y_i \downarrow -\infty$ keeping the rest of the coordinates fixed. The resulting sequence is still in C. However, from (4.69), it follows that the inequality gets reversed unless p_i is non-negative. Since this is true for every $i = 1, 2, \ldots, k+1$, and $p \neq 0$, we can normalise p to obtain

$$r_i \triangleq p_i \Big/ \sum_{i=1}^{k+1} p_j, \quad \forall\, i = 1, 2, \ldots, m+n.$$

The vector r is clearly in $\mathcal{S}_{k+1}$, as defined in (4.60).

Now, recall from (4.57) that $T = -F(\Phi)$, where the function F is given by (4.56). From (4.67), $x_0 \in T$, therefore, there exists a function $\phi^* \in \Phi$ such that

$$x_0 = -F(\phi^*) = \left(-\int \phi^* g_1 \mathrm{d}\mu, -\int \phi^* g_2 \mathrm{d}\mu, \ldots, -\int \phi^* g_{k+1} \mathrm{d}\mu \right).$$

In other words, we can write the second inequality in (4.69) as

$$\langle\, p\,,\ -F(\phi^*)\,\rangle \leq \langle\, p\,,\ -F(\phi)\,\rangle\,, \quad \forall\, \phi \in \Phi.$$

If we expand the function $F(.)$, and interchange summation with integration, we get

$$-\int \left(\sum r_i g_i \right) \phi^* \mathrm{d}\mu \leq -\int \left(\sum r_i g_i \right) \phi \mathrm{d}\mu, \quad \forall\, \phi \in \Phi.$$

This is possible only if

$$\phi^* = \begin{cases} 1, \text{if} \quad \sum_{i=1}^{k+1} r_i g_i(\omega) > 0, \\ \\ 0, \text{if} \quad \sum_{i=1}^{k+1} r_i g_i(\omega) < 0. \end{cases} \tag{4.70}$$

Finally, recall from (4.67) that $x_0 \in Q_\nu$. From (4.65), we infer that

$$x_0(i) \leq -\gamma_i \quad \Leftrightarrow \int \phi^* g_i \mathrm{d}\mu \geq \gamma_i, \quad \forall\, i = 1, 2, \ldots, k.$$

This shows that $\phi^* \in \mathcal{A}$. Also, from (4.65), we get

$$x_0 \leq -d^* \Leftrightarrow \int \phi g_{k+1} \mathrm{d}\mu \geq d^*.$$

However, d^* denotes the maximum value that the function $\mathcal{T}$, of (4.59), achieves on $\mathcal{A}$. Thus clearly

$$\int \phi g_{k+1} \mathrm{d}\mu = d^*$$

and ϕ^*, as in (4.70), maximizes $\mathcal{T}$ over $\mathcal{A}$. This prove the proposition. □

4.5.2 Uniform law of large numbers

We briefly mention here three basic theorems about the theory of uniform law of large numbers and the related concept of Vapnik-Červonenkis dimensions. This is a subject in itself and we shall use very little of it for our purpose. Hence we shall skip all details and refer the reader to the excellent book [Devroye et al., 1996, Chap. 12], from where our propositions in this section have been lifted.

Notation 4.1. We consider a probability space $(\Theta, \Im, \varrho)$, where Θ is a complete, separable metric space. On Θ^n, let ϱ^n denote the product probability measure on the product σ-algebra. Similarly on $\Theta^\infty := \Theta^{\mathbb{N}}$, let ϱ^∞ denote the infinite product probability. For any $\theta \in \Theta^\infty$, and any $n \in \mathbb{N}$, define the random *empirical measure*: $\varrho_n(C) := 1/n \sum_{i=1}^n 1_{(\theta_i \in C)}, \quad C \in \Im$, or, for any $\Im$-integrable function f, the corresponding random expectation $\varrho_n(f) := 1/n \sum_{i=1}^n f(\theta_i)$.

For any $C \in \Im$ and any $\epsilon > 0$, the law of large numbers dictate

$$\lim_{n\to\infty} \varrho^\infty \Big(|\varrho_n(C) - \varrho(C)| > \epsilon \Big) = 0. \tag{4.71a}$$

However, if we have a collection of $\{C_\alpha\}_{\alpha\in I}$ of sets in $\Im$, it is not always true that

$$\lim_{n\to\infty} \varrho^\infty \Big(\sup_{\alpha\in I} |\varrho_n(C_\alpha) - \varrho(C_\alpha)| > \epsilon \Big) = 0. \tag{4.71b}$$

Equality above can be achieved under proper conditions on the collection $\{C_\alpha\}_{\alpha\in I}$, and then we say *Uniform Law of Large Numbers*(ULLN) holds. The Vapnik-

Červonenkis theory provides one such condition. Its strength lies in that the condition on $\{C_\alpha\}_{\alpha\in I}$ is combinatorial in nature, and hence independent from the choice of ϱ. (This sometimes can also be a weakness, since significant improvements can be made for specific choice of ϱ.) The theory begins with the concept of *shatter-coefficient*.

Definition 4.1. Let $\{C_\alpha\}_{\alpha\in I}$ be a collection of $\Im$-measurable subsets of Θ. For $(\theta_1,\ldots,\theta_d)\in\Theta^d$, let $\mathcal{N}(\theta_1,\ldots,\theta_d)$ be the number of different sets in

$$\left\{\quad \{\theta_1,\ldots,\theta_d\}\cap C_\alpha,\ \alpha\in I \quad\right\}.$$

The d-th shatter coefficient of the collection $\{C_\alpha\}_{\alpha\in I}$ is defined as

$$s_d \triangleq \max_{(\theta_1,\ldots,\theta_d)\in\Theta^d} \mathcal{N}(\theta_1,\ldots,\theta_d).$$

In other words, the shatter coefficient is the maximal number of different subsets of d points that can be picked out by the class $\{C_\alpha\}_{\alpha\in I}$.

Remark. Note that we have deliberately suppressed mentioning the class $\{C_\alpha\}_{\alpha\in I}$ in the notation for the shatter coefficient. This is really for notational clarity. The shatter coefficient is clearly a property of the collection of sets we consider.

The following theorem can be found in [Devroye et al., 1996, Thm 12.5, p. 197].

Theorem 4.4. *For any collection $\{C_\alpha\}_{\alpha\in I}$, and for any $n\in\mathbb{N}$, $\epsilon>0$, we have*

$$\varrho^\infty\left\{\sup_{\alpha\in I}|\varrho_n(C_\alpha)-\varrho(C_\alpha)\,|>\epsilon\right\}\le 8s_n\exp(-n\epsilon^2/32), \tag{4.72}$$

where the constant s_n is the nth shatter coefficient of the collection $\{C_\alpha\}_{\alpha\in I}$ and is independent of the probability measure ϱ.

Hence (4.71b) will hold if the constants s_n grows at most polynomially. This is achieved for certain collections of sets which have a finite Vapnik-Čeronenkis (VC) dimension. The following definition is from [Devroye et al., 1996, p. 196].

Definition 4.2. As before we consider the collection $\{C_\alpha\}_{\alpha\in I}$ of $\Im$-measurable subsets of Θ. The largest positive integer for which $s_d = 2^d$ is known as the VC dimension of the collection $\{C_\alpha\}_{\alpha\in I}$. If $s_d = 2^d$ for all integers $d \geq 1$, we then define the VC dimension to be ∞.

The next lemma [Devroye et al., 1996, p. 218] describes a fundamental relationship between VC dimension and the shatter coefficients.

Sauer's Lemma. *Let $\{C_\alpha\}_{\alpha\in I}$ be a subset of $\Im$ with finite VC dimension $\mathcal{V} > 2$. Then for all $n > 2\mathcal{V}$, we have $s_n \leq n^{\mathcal{V}}$.*

Thus Theorem 4.4 together with Sauer's Leamma will yield the following.

Theorem 4.5. *Let $(\Theta, \Im)$ be a measurable space. Let $\{C_\alpha\}_{\alpha\in I}$ be any collection of measurable subsets of Θ with a finite VC dimension $\mathcal{V}$. Then for any probability measure ϱ on $(\Theta, \Im)$ and any $n \geq 2\mathcal{V}$, we have*

$$\varrho^\infty \left\{ \sup_{\alpha\in I} |\varrho_n(C_\alpha) - \varrho(C_\alpha)\ | > \epsilon \right\} \leq 8n^{\mathcal{V}} \exp(-n\epsilon^2/32). \tag{4.73}$$

In particular, $\lim_{n\to\infty} \varrho^\infty \{\sup_{\alpha\in I} |\varrho_n(C_\alpha) - \varrho(C_\alpha)\ | > \epsilon\} = 0$.

We shall be using, however, a better bound than Theorem 4.5. The following theorem is from Devroye (1982).

Theorem 4.6. *In the setting of the previous theorem 4.5, we have*

$$\varrho^\infty \left\{ \sup_{\alpha\in I} |\varrho_n(C_\alpha) - \varrho(C_\alpha)\ | > \epsilon \right\} \leq 4s_{n^2} \exp(-2n\epsilon^2 + 4\epsilon + 4\epsilon^2).$$

Hence, by Sauer's Lemma,

$$\varrho^\infty \left\{ \sup_{\alpha\in I} |\varrho_n(C_\alpha) - \varrho(C_\alpha)\ | > \epsilon \right\} \leq 4n^{2\mathcal{V}_C} \exp(-2n\epsilon^2 + 4\epsilon + 4\epsilon^2). \tag{4.74}$$

Finally There is one collection of sets with finite VC dimensions which we shall need in our analysis.

Proposition 4.3. *[Dudley, 1978, Thm 7.2] Let G be a d-dimensional real vector space of real functions on an infinite set X. Define the class of sets*

$$\mathcal{C} = \{\{x \in X : g(x) > 0\} : g \in G\}.$$

Then the VC dimension of $\mathcal{C}$ is not more than d.

Conclusion

The results in the previous chapters, although general, nevertheless falls short in some examples. The results in Chapter 3 do not hold for an important class of risk measures called the entropy risk, which arises naturally with exponential loss functions (see example 2.1). It is unclear how to proceed in this case when no natural Banach space seems to help. Chapter 4 also has some serious shortcomings. One, that the computations increase exponentially as T increases in a T-period model. The estimates in Theorem 4.6 then has to be improved, and this might already have been done in the machine learning literature, although we are not aware of it. It is worth mentioning, that bounds such as in theorem 4.6 is conservative. Actual simulations converge much faster than predicted. Theoretically, our proofs can be extended to the case where we have infinitely many scenarios, with uniformly integrable Radon-Nikodým derivatives, in the description of the risk measure in (4.1). However the computations become unfeasible. Since our main stress was on computing the strategies, results like Proposition 4.1 is not unnecessarily generalized. More serious is the case of continuous times, which is completely beyond our methodologies. And it will be an interesting project to consider extensions of the results in continuous time. Finally, this thesis is solely concerned with static risk measures, which, although surprisingly flexible to be moulded towards a variety of applications, has its own limitations. A much more theoretically and conceptually difficult concept is that of dynamic risk measures, an excellent source for which is [Klöppel and Scweizer, 2005]. Much of this area is still left for exploring and will be a likely direction for future work.

Bibliography

[Artzner et al., 1999] Artzner, P., Delbaen, F., Eber, J., and Heath, D. (1999). Coherent measures of risk. *Math. Finance*, 9:203–228.

[Barrieu and El Karoui, 2005a] Barrieu, P. and El Karoui, N. (2005a). Inf-convolution of risk measures and optimal risk transfer. *Finance & Stochastics*, 9:269–298.

[Barrieu and El Karoui, 2005b] Barrieu, P. and El Karoui, N. (2005b). Pricing, hedging and optimally designing derivatives via minimization of risk measures. In *To appear in Volume on Indifference Pricing*. Princeton University Press.

[Carr et al., 2001] Carr, P., Geman, H., and Madan, D. (2001). Pricing and hedging in incomplete markets. *J. Financial Economics*, 62:131–167.

[Černy and Hodges, 2001] Černy, A. and Hodges, S. (2001). The theory of good-deal pricing in incomplete markets. In *Mathematical Finance - Bachelier Congress 2000*, pages 175–202. Springer-Verlag, Berlin.

[Cherny, 2005] Cherny, A. (2005). Pricing with coherent risk. Unpublished. Downloadable from http://mech.math.msu.su/ cherny/poem1.pdf.

[Cochrane and Saá-Requejo, 2000] Cochrane, J. H. and Saá-Requejo, J. (2000). Beyond arbitrage: Good-deal asset price bounds in incomplete markets. *Journal of political economy*, 108(1):79–119.

[Cvitanic and Karatzas, 1992] Cvitanic, J. and Karatzas, I. (1992). Convex duality in constrained portfolio optimization. *Annals of Applied Probability*, 2(4):767–818.

[Cvitanic and Karatzas, 1993] Cvitanic, J. and Karatzas, I. (1993). Hedging contingent claims with constrained portfolios. *Annals of Applied Probability*, 3(3):652–681.

[Devroye et al., 1996] Devroye, L., Györfi, L., and Lugosi, G. (1996). *A probabilistic Theory of Pattern Recognition.*, volume 31 of *Applications of Mathematics*. Springer-Verlag, New York.

[Dudley, 1978] Dudley, R. M. (1978). Central limit theorems for empirical measures. *The Annals of Probability*, 6(6):899–929.

[El Karoui and Quenez, 1995] El Karoui, N. and Quenez, M. C. (1995). Dynamic programming and pricing of contingent claims in an incomplete market. *SIAM Journal on Control and Optimization*, 33(1):29–66.

[Ferguson, 1967] Ferguson, T. (1967). *Mathematical Statistics: a decision theoretic approach*. Academic Press.

[Föllmer and Leukert, 1999] Föllmer, H. and Leukert, P. (1999). Quantile hedging. *FS*, 3:251–273.

[Föllmer and Leukert, 2000] Föllmer, H. and Leukert, P. (2000). Efficient hedging: cost versus shortfall risk. *FS*, 4:117–146.

[Föllmer and Schied, 2002] Föllmer, H. and Schied, A. (2002). Convex measures of risk and trading constraints. *Finance & Stochastics*, 6:429–447.

[Föllmer and Schied, 2004] Föllmer, H. and Schied, A. (2004). *Stochastic Finance: An Introduction in Discrete Time*, volume 27 of *Studies in Mathematics*. de Gruyter, Berlin., second edition.

[Gelman et al., 2003] Gelman, A., Carlin, J. B., Stern, H. S., and Rubin, D. B. (2003). *Bayesian Data Analysis*. Chapman & Hall/CRC, London, second edition.

[Goebel and Reich, 1984] Goebel, K. and Reich, S. (1984). *Uniform Convexity, Hyperbolic Geometry, and Nonexpansive Mappings.*, volume 83 of *Pure and Applied Mathematics*. Marcel Dekker, Inc., New York.

[Heath, 2000] Heath, D. (2000). Back to the future. In *Plenary lecture at the First World Congress of the Bachelier Society*, Paris.

[Huber, 1981] Huber, P. (1981). *Robust Statistics*. Wiley, New York.

[Jaschke and Küchler, 2001] Jaschke, S. and Küchler, U. (2001). Coherent risk measures and good-deal bounds. *Finance & Stochastics*, 5:181–200.

[Karatzas, 1996] Karatzas, I. (1996). *Lectures on the Mathematics of Finance*, volume 8 of *CRM monograph series*. American Mathematical Society.

[Karatzas and Shreve, 1991] Karatzas, I. and Shreve, S. E. (1991). *Brownian Motion and Stochastic Calculus.*, volume 113 of *Graduate Texts in Mathematics*. Springer-Verlag, New York., second edition.

[Klöppel and Scweizer, 2005] Klöppel, S. and Scweizer, M. (2005). Dynamic utility indifference valuation via convex risk measures. Working paper No. 209, NCCR: Financial Valuation and Risk Managaement.

[Larsen et al., 2004] Larsen, K., Pirvu, T., Shreve, S., and Tütüncü, R. (2004). Satisfying convex risk limits by trading. *Finance & Stochastics*, 9:177–195.

[Megginson, 1998] Megginson, R. E. (1998). *An Introduction to Banach Space Theory*, volume 183 of *Graduate Texts in Mathematics*. Springer-Verlag.

[Protter, 2004] Protter, P. (2004). *Stochastic Integration and Differential Equations*, volume 21 of *Stochastic Modelling and Applied Probability*, pages 244–245. Springer-Verlag., second edition.

[Rockafellar, 1997] Rockafellar, R. (1997). *Convex Analysis*. Princeton Landmarks in Mathematics series. Princeton University Press, tenth printing and first paperback printing edition.

[Rudin, 1991] Rudin, W. (1991). *Functional Analysis*. International Series in Pure and Applied Mathematics. McGraw-Hill, second edition.

[Schachermayer, 2000] Schachermayer, W. (2000). Optimal investment in incomplete financial markets. In *Proceedings of the First World Congress of the Bachelier Society*. Springer.

[Staum, 2004] Staum, J. (2004). Fundamental theorems of asset pricing for good deal bounds. *Mathematical Finance*, 14(2):141–161.