Introduction to convex optimization in financial markets

Teemu Pennanen*

April 11, 2012

Abstract

Convexity arises quite naturally in financial risk management. In risk preferences concerning random cash-flows, convexity corresponds to the fundamental diversification principle. Convexity is a basic property also of budget constraints both in classical linear models as well as in more realistic models with transaction costs and constraints. Moreover, modern securities markets are based on trading protocols that result in convex trading costs. The first part of this paper gives an introduction to certain basic concepts and principles of financial risk management in simple optimization terms. The second part reviews some convex optimization techniques used in mathematical and numerical analysis of financial optimization problems.

1 Introduction

Financial risks can be managed by trading in financial markets. By appropriate trading, individuals and financial institutions may modify their net cash-flows to better conform to their risk preferences. For example, a home owner may be able to achieve a more attractive risk profile for his future cash-flows by buying a home insurance. Insurers, on the other hand, invest insurance premiums in financial markets in order to optimize their net cash-flow structure resulting from delivering insurance claims and collecting investment returns. The same principle is behind the classical Black–Scholes–Merton option pricing framework, where the seller of an option invests the premium in financial markets according to an investment strategy whose return matches the option payout.

Traditionally, the main tools in mathematical finance have come from stochastics but convex analysis is turning out to be equally useful. Techniques of convex analysis allow for extending traditional models of financial markets to more realistic ones with e.g. market frictions and constraints that are often encountered in practical applications. Moreover, the optimization perspective brings in variational and computational techniques that have been successful in more

*Department of Mathematics, King's College London, teemu.pennanen@kcl.ac.uk

traditional fields of applied mathematics such as partial differential equations and operations research.

This paper gives a brief introduction to financial risk management from the point of view of convex optimization. After a description of market models and some basic problems in financial risk management, we give a brief review of convex optimization techniques used in the analysis of such problems. We use simple discrete time models which do not demand prior knowledge of stochastic analysis beyond basic measure theory. In fact, many important issues can be introduced already in a single-period setting where many risk management problems reduce to finite-dimensional optimization problems. However, we do allow for general probability spaces since they are needed in most models of financial data. In dynamic settings, this results in infinite-dimensional optimization problems. Certain aspects of convex optimization in markets models with finite probability spaces are treated in Pliska [64] and King [40].

Convexity is often indispensable in mathematical and numerical analysis of financial optimization problems. For example, general characterizations of the *no-arbitrage* property of a perfectly liquid market model in terms of *martingale measures* is largely based on separation theorems for convex sets; see Föllmer and Schied [25] and Delbaen and Schachermayer [16] for comprehensive study of the classical linear model of financial markets. Techniques of convex optimization allow also for significant generalizations of the linear model. Models with transaction costs and portfolio constraints have been studied e.g. in Davis and Norman [15], Dermody and Prisman [18], Dermody and Rockafellar [19, 20], Cvitanic and Karatzas [12, 13], Jouini and Kallal [35, 34], Kabanov [38], Föllmer and Schied [24], Evstigneev, Schürger and Taksar [23], Dempster, Evstigneev and Taksar [17], Schachermayer [71]. An extensive study of models with proportional transaction costs can be found in Kabanov and Safarian [39]. The present paper builds largely on the discrete-time market model introduced in Pennanen [57], where convex analysis and the theory of normal integrands were taken as the main tools of analysis. The model provides a flexible framework for modeling portfolio constraints and transaction costs as well as certain illiquidity effects that arise in modern securities markets.

Besides mathematical analysis, convex optimization offers computational possibilities for risk management beyond the techniques of stochastic analysis alone. Realistic models of risk management often require a combination of techniques from stochastics and optimization. The presence of stochastic elements in a model results in difficult, often infinite-dimensional optimization problems that require specialized optimization techniques where, again, convexity plays a key role.

This paper is organized as follows. Section 2 introduces certain basic problems in financial risk management using a simple one-period model of liquid financial markets. Illiquidity effects arising in modern limit order markets is discussed in Section 2.4. Section 3 gives dynamic extensions in a discrete time setting. Section 4 gives an overview of convex duality in financial models. This part illustrates the role of conjugate duality, theory of normal integrands and recession analysis in the analysis of nonlinear market models. Section 5 outlines briefly some numerical techniques for stochastic optimization. The only novel part of the paper is contained in Section 5.2, which gives an upper bound on the information based complexity of static stochastic optimization problems.

2 Static models

Consider a financial market where a finite set J of assets can be traded. In this section, we will consider *static* problems, where assets are traded only at two dates, the present time $t = 0$ and some time $t = 1$ in the future. Moreover, we assume that the market is perfectly liquid so that the unit prices of the traded assets do not depend on our actions. The unit price of asset $j \in J$ at time t will be denoted by s_t^j. We will assume that the vector $s_0 = (s_0^j)_{j\in J}$ of current prices is known to us before we trade at time $t = 0$. The price vector $s_1 = (s_1^j)_{j\in J}$ will remain uncertain until time $t = 1$. We will model s_1 as a random vector on a probability space $(\Omega, \mathcal{F}, P)$. That is, s_1 is an $\mathcal{F}$- measurable function from Ω to $\mathbb{R}^J$. The linear space of equivalence classes of real-valued random variables will be denoted by $L^0(\Omega, \mathcal{F}, P)$.

Buying a portfolio $x = (x^j)_{j\in J}$ of assets at time $t = 0$ costs $s_0 \cdot x$ units of cash. If we hold on to the portfolio, it will be worth $s_1 \cdot x$ at time $t = 1$. Clearly, $s_1 \cdot x$ is random. Here x^j denotes the number of units of asset $j \in J$ we hold. It requires investing $h^j = s_0^j x^j$ units of cash in asset j at time $t = 0$. The value of our portfolio at time $t = 1$ can be expressed $r \cdot h$, where $r = (r^j)_{j\in J}$ is the random vector with components $r^j = s_1^j/s_0^j$. While in practice, it is usually more common to describe investments in terms of cash invested in each asset, formulations in terms of units are sometimes more convenient in mathematical analysis.

2.1 Asset-liability management

Consider the problem

$$\begin{array}{llll} \text{minimize} & \mathcal{V}(c - s_1 \cdot x) & \text{over} & x \in \mathbb{R}^J, \\ \text{subject to} & s_0 \cdot x \le w, & x \in D, & \end{array} \tag{ALM}$$

where $w \in \mathbb{R}$ is a given initial wealth, $D \subseteq \mathbb{R}^J$ describes portfolio constraints, $c \in L^0(\Omega, \mathcal{F}, P)$ is a random amount of cash (a *claim*) the agent must deliver at time $t = 1$ and the function $\mathcal{V} : L^0(\Omega, \mathcal{F}, P) \to \overline{\mathbb{R}}$ measures the "risk/disutility/regret" from the random net expenditure $c - s_1 \cdot x$ at time $t = 1$. Problem (ALM) can be viewed as an *asset-liability management* problem where one is trying to find a portfolio x of assets whose value matches the liabilities described by c optimally as measured by $\mathcal{V}$. We allow c to take arbitrary real values so it may describe costs as well as income.

We will assume throughout that D is a convex set containing the origin and that $\mathcal{V}$ is monotonic, normalized and convex. Monotonicity means that $\mathcal{V}(c_1) \le \mathcal{V}(c_2)$ whenever $c_1 \le c_2$ almost surely, i.e. one always prefers more cash. Normalization means that $\mathcal{V}(0) = 0$. It is posed mainly for notational convenience. As long as $0 \in \operatorname{dom} \mathcal{V}$, normalization can be achieved by adding a constant to the objective. Convexity is one of the basic axioms of risk measures; see e.g. Artzner, Delbaen, Eber and Heath [1], Föllmer and Schied [25], Pflug and Römisch [63] or Rockafellar [68] for general discussions on quantitative risk measurement and description of risk preferences. Here, however, the values of $\mathcal{V}$ need not reflect the present values of uncertain cash-flows. Such values will be studied in Sections 2.2 and 2.3 below.

The classical utility maximization problem corresponds to $\mathcal{V}(c) = -Eu(-c)$, where E denotes the expectation and u is a concave nondecreasing function on the real line.[1] When $\mathcal{V}(c) = \inf_\alpha E[c + \lambda|c - \alpha|^2]$ for a positive scalar λ, we recover the classical mean-variance criterion of Markowitz [46]. It should be noted, however, that the mean-variance criterion is not monotonic. One could also take

$$\mathcal{V}(c) = \inf_{\alpha\in\mathbb{R}} \{\alpha + Ev(c - \alpha)\}$$

for a convex function v on $\mathbb{R}$ as proposed by Ben-Tal and Teboulle [5, 6]. This covers the mean-variance formulation as well as the Conditional Value at Risk (choose $v(c) = c+\lambda|c|^2$ and $v(c) = \max\{c, 0\}/(1-\gamma)$, respectively); see Rockafellar and Uryasev [67, 68]. The greatest among all convex monotonic normalized functions on L^0 is[2]

$$\mathcal{V} = \delta_{L^0_-}, \tag{1}$$

where $L^0_- := \{c \in L^0(\Omega, \mathcal{F}, P) \,|\, c \le 0 \ P\text{-almost surely}\}$. This corresponds to a completely risk averse agent who deems all losses unacceptable. Note that (1) can be expressed as $\mathcal{V}(c) = -Eu(-c)$ with $u = \delta_{\mathbb{R}_+}$.

Problem (ALM) depends essentially on the agent's subjective views and preferences described by the probability distribution P and the function $\mathcal{V}$, respectively. The optimum value and solutions of (ALM) depend also on the agent's financial position as described by the initial wealth w and the liabilities c. In pricing and hedging of claims as well as in determining capital requirements for financial liabilities one is concerned with how the risk profile depends on a given financial position. We will denote the optimum value of (ALM) by

$$\varphi(w, c) := \inf\,(\text{ALM})$$

The convexity of D and $\mathcal{V}$ imply that (ALM) is a convex optimization problem. Indeed, the objective is the composition of $\mathcal{V}$ with the linear function $x \mapsto s_1 \cdot x$ from $\mathbb{R}^J$ to $L^0(\Omega, \mathcal{F}, P)$ so it is convex as a function of x. In the most risk averse case with (1), we have

$$\varphi(w, c) = \delta_{\mathcal{C}}(w, c),$$

[1] Throughout, we define the expectation of a random variable as $+\infty$ unless its positive part has a finite expectation. The expectation is then well-defined for any random variable.

[2] Here and in what follows δ_C denotes the *indicator function* of a set C: $\delta_C(x) = 0$ if $x \in C$ and $\delta_C(x) = +\infty$ otherwise.

where

$$\mathcal{C} = \{(w,c) \in \mathbb{R} \times L^0 \,|\, \exists x \in D : \; s_0 \cdot x \le w, \; c \le s_1 \cdot x \; P\text{-a.s.}\}.$$

On the other hand, the value function can always be written in terms of $\mathcal{C}$ as

$$\varphi(w,c) = \inf_d \{\mathcal{V}(c-d) \,|\, (w,d) \in \mathcal{C}\} = \inf_d \{\mathcal{V}(d) \,|\, (w, c-d) \in \mathcal{C}\}.$$

The following simple fact turns out to have important consequences in financial risk management.

Lemma 1 *The value function φ is convex.*

Proof. We have $\varphi(w,c) = \inf_d F(w,c,d)$, where $F(w,c,d) = \mathcal{V}(d) + \delta_{\mathcal{C}}(w, c-d)$. The convexity of $\mathcal{V}$ and $\mathcal{C}$ imply that of F, which in turn implies the convexity of φ; see e.g. [66, Theorem 1]. $\square$

2.2 Capital requirements

In risk measurement, accounting, financial reporting and supervision of financial institutions, one is often interested in determining the least amount of capital that would suffice for "covering" a financial liability. In the one-period model of Section 2.1, where financial liabilities are described by random cash-flows $c \in L^0(\Omega, \mathcal{F}, P)$ at time $t = 1$, such an amount can be defined as the optimum value of the problem

$$\begin{array}{lll} \text{minimize} & w \quad \text{over} & w \in \mathbb{R}, \; x \in \mathbb{R}^J, \\ \text{subject to} & s_0 \cdot x \le w, \quad x \in D, & \\ & \mathcal{V}(c - s_1 \cdot x) \le 0. & \end{array} \tag{2}$$

The optimum value $\pi_0(c)$ gives the least amount of initial capital one needs in order to construct a *hedging strategy* (a portfolio) x whose value at time $t = 1$ covers a sufficient part of the claim so that the risk associated with the residual is no higher than the risk from doing nothing at all (recall that $\mathcal{V}$ is normalized: $\mathcal{V}(0) = 0$). The capital requirement $\pi_0(c)$ defines an extended-real valued function on the space $L^0(\Omega, \mathcal{F}, P)$ of cash-flows at time $t = 1$.

In the completely risk averse case with (1), the capital requirement π_0 coincides with the well-known *superhedging cost*

$$\pi_{\sup}(c) := \inf\{w \,|\, \exists x \in D : s_0 \cdot x \le w, \; c \le s_1 \cdot x \; P\text{-a.s.}\}; \tag{3}$$

see [25, 16, 39] and their references. Since (1) is the greatest among all convex monotonic normalized functions on $L^0(\Omega, \mathcal{F}, P)$, we always have $\pi_0 \le \pi_{\sup}$. On the other hand, if we ignore the existence of financial markets and assume that all wealth is invested "under the mattress", we get

$$\pi_0(c) = \inf\{w \,|\, \mathcal{V}(c - w) \le 0\}$$

which corresponds to classical *premium principles* from actuarial mathematics; see e.g. Bühlmann [9]. By an appropriate choice of a trading strategy, one may be able to lower the capital requirement. The same principle is behind the famous Black–Scholes–Merton option pricing model [8] where the price of an option is defined as the least amount of initial capital that allows for the construction of a trading strategy whose terminal value equals the payout of the option. Unlike in the Black–Scholes–Merton model, however, exact replication is often impossible in practice so one has to define preferences concerning the uncertain amount of wealth remaining after the delivery of the claim.

Our formulation of (2) is also motivated by modern financial supervisory standards, such as the Solvency II Directive 2009/138/EC of the European Parliament, which promotes market consistent accounting principles that recognize the risks in both assets and liabilities. The interplay between capital requirements and asset management was recently studied in Artzner, Delbaen and Koch-Medona [2]. An implementation of a dynamic version of (2) (see Section 3.2 below) to the valuation of pension insurance liabilities is described in Hilli, Koivu and Pennanen [30].

Under mild conditions, the capital requirement $\pi_0(c)$ can be expressed in terms of the value function φ of (ALM). This turns out to be useful both in duality theory as well as in numerical computations.

Proposition 2 *The function π_0 is convex, monotonic and $\pi_0(0) \leq 0$. Moreover,*

$$\pi_0(c) = \inf\{w \mid \varphi(w,c) \leq 0\}$$

if either of the following conditions hold:

(a) The optimum in (ALM) *is attained for every w and c.*

(b) $\varphi(w,c)$ is strictly decreasing in w.

Proof. The convexity is again a simple application of [66, Theorem 1] and monotonicity is immediate while the normalization property follows from that of $\mathcal{V}$ by choosing $x = 0$ in (2). If $w > \pi_0(c)$, there exists, by definition of π_0, an $x \in D$ such that $s_0 \cdot x \leq w$ and $\mathcal{V}(c - s_1 \cdot x) \leq 0$. This implies $\varphi(w,c) \leq 0$ so we must have $\pi_0(c) \geq \inf\{w \mid \varphi(w,c) \leq 0\}$. To prove the converse, let $w \in \mathbb{R}$ be such that $\varphi(w,c) \leq 0$. It suffices to show that $\pi_0(c) \leq w$. Under condition (a), there is an $x \in D$ such that $s_0 \cdot x \leq w$ and $\mathcal{V}(c - s_1 \cdot x) \leq 0$. Thus, $\pi_0(c) \leq w$. Under condition (b), we have for every $w' > w$ that $\varphi(w',c) < 0$ so there is an $x \in D$ such that $s_0 \cdot x \leq w'$ and $\mathcal{V}(c - s_1 \cdot x) \leq 0$. Thus, $\pi_0(c) \leq w'$. Since $w' > w$ was arbitrary, we get $\pi_0(c) \leq w$. $\square$

The capital requirement π_0 may be interpreted much like *risk measures* in [1]. In general, however, π_0 lacks the "translation-invariance" property that $\pi_0(c + \alpha) = \pi_0(c) + \alpha$ for all $\alpha \in \mathbb{R}$ which is often required of risk measures; see e.g. [25]. While convexity and monotonicity of $\mathcal{V}$ imply the convexity and monotonicity of π_0, the same does not hold for cash-invariance. However, as long as long the market model allows for long positions in cash, we have $\pi_0(c+\alpha) \leq$

$\pi_0(c)+\alpha$ for all $\alpha \geq 0$ as is easily verified. Such a "cash sub-additivity" property has been recently studied in El Karoui and Ravanelli [22] and Drapeau and Kupper [21]. The full "translation-invariance" property holds if arbitrary long and short positions in cash are allowed. While this is a standard assumption in financial mathematics (see Example 4 below), it rarely holds in practice.

Condition (b) in Proposition 2 is quite natural as it means that an increase in the initial endowment always leads to a strictly preferred net position at the end of the holding period. A sufficient condition for (a) to hold will be given in Theorem 6 below.

In general, the value of $\pi_0(c)$ depends on the the risk preferences described by $\mathcal{V}$ and the probability measure P describing the future development of the financial markets and the liabilities. These are both subjective factors. In supervisory frameworks, they should be specified according to the supervisor's views.

When (2) cannot be solved exactly, the capital requirement should be defined as the least value a financial institution *can* achieve in (2). It then depends also on the trading expertise of a financial institution. Institutions that are good at hedging their liabilities can operate at lower capital.

2.3 Pricing of contingent claims

Consider now the problem of valuing a contingent claim $c \in L^0(\Omega, \mathcal{F}, P)$ from the point of view of an agent whose current financial position is given by an initial wealth $\bar{w} \in \mathbb{R}$ and liabilities $\bar{c} \in L^0(\Omega, \mathcal{F}, P)$. This time we are not looking for the amount of wealth that allows the agent to survive with a given liability but a price at which he would be willing to sell a claim c. It is intuitively clear that such a price should depend on the agent's financial position. The financial position matters also when buying a claim. For instance, the value of a European option to an agent probably depends on whether the agent owns the underlying stock. Similarly, a wheat farmer is more likely to buy a futures contract on wheat than somebody who's income does not depend on wheat price. In fact, most financial instruments exist because of the differences between financial positions of different agents.

The least cash-payment at time $t = 0$ that would suffice as a compensation for delivering c at time $t = 1$ can be expressed in the model of Section 2.1 as

$$\pi(\bar{w}, \bar{c}; c) = \inf\{w \mid \varphi(\bar{w} + w, \bar{c} + c) \leq \varphi(\bar{w}, \bar{c})\}, \tag{4}$$

where again, $\varphi(w, c)$ denotes the optimum value of (ALM). Any price $w > \pi(\bar{w}, \bar{c}; c)$ would allow the agent to optimize his portfolio so that the risk with the new financial position $(\bar{w} + w, \bar{c} + c)$ is no greater than it would be without the trade. This kind of *indifference principles* in pricing of contingent claims go back at least to Hodges and Neuberger [32]; see Carmona [10] for a recent account. It is easily checked that the convexity of φ implies that $\pi(\bar{w}, \bar{c}; \cdot)$ is a convex function on $L^0(\Omega, \mathcal{F}, P)$.

While π gives the least price an agent would be willing to sell a claim for, the greatest buying price for c is given by

$$\pi^b(\bar{w},\bar{c};c) = \sup\{w \,|\, \varphi(\bar{w}-w,\bar{c}-c) \le \varphi(\bar{w},\bar{c})\}.$$

Clearly, $\pi^b(\bar{w},\bar{c};c) = -\pi(\bar{w},\bar{c};-c)$. When there are no constraints, i.e. when $D = \mathbb{R}^J$, we can bound the selling and buying prices of a contingent claim between the superhedging cost $\pi_{\sup}(c)$ given by (3) and the "subhedging cost"

$$\pi_{\inf}(c) := \sup\{w \,|\, \exists x \in D : s_0 \cdot x + w \le 0,\ c + s_1 \cdot x \ge 0\ P\text{-a.s.}\}.$$

Clearly, $\pi_{\inf}(c) = -\pi_{\sup}(-c)$. A claim $c \in L^0(\Omega,\mathcal{F},P)$ is said to be *replicable* (or attainable; see e.g. [25]) if there is an $x \in \mathbb{R}^J$ such that $s_1 \cdot x = c$ almost surely.

Proposition 3 *If $\pi(\bar{w},\bar{c};0) \ge 0$, then*

$$\pi^b(\bar{w},\bar{c};c) \le \pi(\bar{w},\bar{c};c).$$

If there are no portfolio constraints, then

$$\pi(\bar{w},\bar{c};c) \le \pi_{\sup}(c).$$

In particular, when both conditions hold,

$$\pi_{\inf}(c) \le \pi^b(\bar{w},\bar{c};c) \le \pi(\bar{w},\bar{c};c) \le \pi_{\sup}(c)$$

with equalities throughout when c is replicable.

Proof. By convexity of $\pi(\bar{w},\bar{c};\cdot)$,

$$\pi(\bar{w},\bar{c};0) \le \frac{1}{2}\pi(\bar{w},\bar{c};c) + \frac{1}{2}\pi(\bar{w},\bar{c};-c)$$

so $\pi(\bar{w},\bar{c};0) \ge 0$ implies $-\pi(\bar{w},\bar{c};-c) \le \pi(\bar{w},\bar{c};c)$ which is the first claim.

For any $w > \pi_{\sup}(c)$, there is an $x' \in \mathbb{R}^J$ such that $s_0 \cdot x' \le w$ and $c \le s_1 \cdot x'$ almost surely. When there are no constraints, we can make the change of variables $x \to x - x'$ in (ALM) giving

$$\varphi(\bar{w},\bar{c}) = \varphi(\bar{w} + s_0 \cdot x', \bar{c} + s_1 \cdot x').$$

By monotonicity of $\mathcal{V}$,

$$\varphi(\bar{w},\bar{c}) \ge \varphi(\bar{w} + w, \bar{c} + c)$$

so $\pi(\bar{w},\bar{c};c) \le w$. Since $w > \pi_{\sup}(c)$ was arbitrary, we must have $\pi(\bar{w},\bar{c};c) \le \pi_{\sup}(c)$. The last claim now follows from the fact that $\pi^b(\bar{w},\bar{c};c) = -\pi(\bar{w},\bar{c};-c)$ and $\pi_{\inf}(c) = -\pi_{\sup}(-c)$ and by noting that if $s_1 \cdot x = c$, then $\pi_{\sup}(c) \le s_0 \cdot x \le \pi_{\inf}(c)$. $\square$

The condition $\pi(\bar{w}, \bar{c}; 0) \geq 0$ means that one cannot lower the initial wealth without affecting the optimum value of (ALM). It holds, in particular, if $\varphi(w, c)$ is strictly decreasing in the initial endowment w. In general, selling and buying prices depend on an agent's financial position $(\bar{w}, \bar{c})$, future views as described by P and risk preferences $\mathcal{V}$, all of which are subjective factors. The inequality $\pi_b(\bar{c}; c) \leq \pi(\bar{c}; c)$ just means that two agents with identical characteristics have no incentive to trade with each other. By the second part of Proposition 3, prices of replicable claims are given by the superhedging cost $\pi_{\sup}$ which is independent of such subjective factors. Moreover, the convexity of $\pi_{\sup}$ and the concavity of $\pi_{\sup}$ imply that, on the space of replicable claims, prices are linear in c.

The pricing principle (4) assumes that one can solve (ALM) to optimality. In practice, this may be impossible, but (4) still makes sense if one redefines φ as the lowest value one *can* achieve in (ALM). Besides the financial position $(\bar{w}, \bar{c})$, future views P and risk preferences $\mathcal{V}$, offered prices thus depend on an agent's expertise in optimizing his portfolio before and after the trade.

2.4 Illiquidity

The market model considered so far describes perfectly *liquid* markets where the unit price of a security does not depend on whether we are buying or selling nor on the quantity of the traded amount. In reality, different unit prices are associated with buying and selling and, moreover, as the traded quantities increase, the prices start to move against us. This is often referred to a *illiquidity*.

Many securities are traded in *limit order markets*, where market participants submit buying or selling offers characterized by limits on quantity and unit price; see Harris [26] for a general account of various trading protocols. When buying securities, the quantity available at the lowest submitted selling price is finite. When buying more, one gets the second lowest price and so on. The marginal price $c(x)$ of buying is thus a piecewise constant nondecreasing function of the number x of units bought. Thus, the total cost

$$C(x) = \int_0^x c(z)dz$$

of buying x units is a piecewise linear convex function on $\mathbb{R}_+$. Analogously, the marginal price for selling securities is a nonincreasing piecewise constant function of the number of units sold. Thus, the total revenue $R(x)$ of selling x units is a piecewise linear concave function on $\mathbb{R}_+$.

Given the cost and revenue functions C and R, the market is "cleared" by solving the convex optimization problem

$$\text{maximize} \quad R(x) - C(x) \quad \text{over} \quad x \in \mathbb{R}_+.$$

Optimal solutions are characterized by the *market clearing* condition

$$\partial R(x) - \partial C(x) \ni 0,$$

where the set-valued mappings $\partial R : \mathbb{R} \rightrightarrows \mathbb{R}$ and $\partial C : \mathbb{R} \rightrightarrows \mathbb{R}$ are obtained from the marginal price functions by closing the gaps in the graphs by vertical lines. In the terminology of convex analysis, ∂R and ∂C are the *subdifferentials* of R and C; see [65]. When multiple solutions exist, the greatest solution $\bar{x}$ is implemented by matching $\bar{x}$ units of the most generous selling and buying offers. The prices of the remaining selling offers are then all strictly higher than the prices of the remaining buying offers and no more trades are possible before new offers arrive.

The offers remaining after market clearing are recorded in the so called *limit order book*. It gives the marginal prices for buying or selling a given quantity at the best available prices. Interpreting negative purchases as sales, the marginal prices can be incorporated into a single function $x \mapsto s(x)$ giving the marginal price for buying positive or negative quantities of the commodity. Figure 2.4 presents an example of a marginal price curve s taken from Copenhagen stock exchange. Since the highest buying price is lower than the lowest selling price, the marginal price curve s is a nondecreasing piecewise constant function on $\mathbb{R}$. Consequently,

$$S(x) = \int_0^x s(z)dz$$

defines a piecewise linear convex function on $\mathbb{R}$. It gives the cost of buying x shares. Again, negative x is interpreted as sales and a negative cost as revenue. Note that the perfectly liquid market model studied earlier corresponds to marginal prices $s(x)$ being independent of the traded quantity x in which case the cost function S is linear.

Sections 2.1–2.3 could be readily extended to allow for illiquidity effects. We will do this in the next section in a dynamic setting. Another generalization of the classical linear model of financial markets was proposed in Kabanov [38]; see also Kabanov and Safarian [39]. In Kabanov's model, all assets are treated symmetrically, much as in currency markets, and the trading constraints are described in terms of "solvency cones", which can be interpreted as the negatives of the sets of portfolios that are freely available in the market. In the model of limit order markets, such a set can be expressed as $\{x \in \mathbb{R}^J \,|\, S(x) \le 0\}$. In Kabanov's original model the sets were polyhedral cones. Extensions to more general convex sets were studied in Pennanen and Penner [61].

3 Dynamic models

Consider now a dynamic setting where uncertainty is still modeled by a probability space $(\Omega, \mathcal{F}, P)$ but now one may trade at multiple points $t = 0, \dots, T$ in time. The information available at time t is described by a σ-algebra $\mathcal{F}_t \subseteq \mathcal{F}$ in the sense that, at time t, we do not know which scenario ω will eventually realize but only which element of $\mathcal{F}_t$ it belongs to. Assuming perfect memory, we have $\mathcal{F}_t \subseteq \mathcal{F}_{t+1}$, i.e. the sequence $(\mathcal{F}_t)_{t=0}^T$ is a *filtration.*

A portfolio x_t chosen at time t may depend on the information observed so far but not on information that will only be observed in the future. This

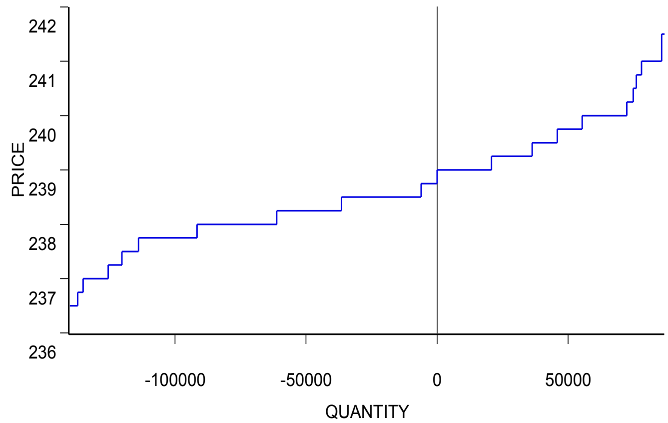

Figure 1: Marginal price curve for shares of the Danish telecom company TDC A/S observed in Copenhagen Stock Exchange on January 12, 2005 at 13:58:19.43. The horizontal axis gives the cumulative depth of the book measured in the number of shares. Negative order quantity corresponds to a sale. The prices are in Danish krone. The data was provided by OMX market research.

means that x_t is an $\mathcal{F}_t$-measurable function from Ω to $\mathbb{R}^J$, or in other words, the trading strategy $x = (x_t)_{t=0}^T$ is *adapted*[3] to the filtration $(\mathcal{F}_t)_{t=0}^T$. The linear space of adapted trading strategies will be denoted by $\mathcal{N}$. Unless $\mathcal{F}_T$ has only a finite number of elements with positive probability, $\mathcal{N}$ is an infinite-dimensional space. We will assume that $\mathcal{F}_0 = \{\Omega, \emptyset\}$ so that x_0 is independent of ω.

The financial market will be described by an $(\mathcal{F}_t)_{t=0}^T$-adapted sequence $S = (S_t)_{t=0}^T$ of *normal integrands* on $\mathbb{R}^J \times \Omega$ such that $S_t(\cdot, \omega)$ is convex and $S_t(0, \omega) = 0$ for each $\omega \in \Omega$. More precisely, S_t is an extended real-valued function on $\mathbb{R}^J \times \Omega$ such that $\omega \mapsto \operatorname{epi} S_t(\cdot, \omega)$ is closed convex-valued and $\mathcal{F}_t$-measurable[4]. Such a sequence S is called a *convex cost process*. The value of $S_t(x, \omega)$ is interpreted as the cost we would have to pay for a portfolio $x \in \mathbb{R}^J$ at time t in state ω; see Section 2.4. The classical linear market model corresponds to $S_t(x, \omega) = s_t(\omega) \cdot x$, where s_t is an $\mathcal{F}_t$-measurable $\mathbb{R}^J$-valued function; see [25] and [16] for thorough treatment for such models. Proportional transaction costs as well as bid-ask-spreads can be represented by sublinear cost functions; see e.g. [35, 13]. More general convex cost processes have been proposed in Çetin and Rogers [11] and Malo and Pennanen [45].

[3]Some authors (e.g. [25, 16, 39]) describe trading strategies by "predictable" processes $H = (H_t)_{t=0}^T$ with H_t denoting the portfolio that was chosen at time $t-1$. This is only a notational difference with $H_t = x_{t-1}$.

[4]A set-valued mapping $\omega \mapsto C(\omega)$ is $\mathcal{F}_t$-measurable if $\{\omega \in \Omega \,|\, C(\omega) \cap U \neq \emptyset\} \in \mathcal{F}_t$ for every open set U.

The measurability condition implies that if $x_t : \Omega \to \mathbb{R}^J$ is $\mathcal{F}_t$-measurable, then $\omega \mapsto S_t(x_t(\omega), \omega)$ is $\mathcal{F}_t$-measurable as well; see [69, Corollary 14.34]. If we are endowed with a random sequence $(d_t)_{t=0}^T$ of cash-flows, then our budget constraint can be written as

$$S_t(\Delta x_t(\omega), \omega) \leq d_t(\omega) \quad P\text{-a.s. } t = 0, \dots, T,$$

where $\Delta x_t := x_t - x_{t-1}$. The interpretation is that the cost of updating the portfolio from x_{t-1} to x_t cannot be more than the amount of cash d_t available at time t.

We do not assume a priori that cash can be transferred freely in time. While positive amounts of cash can usually be transferred easily, the same does not apply to negative amounts, i.e. borrowing is usually more restricted that lending. We can, however, include cash in the set J of traded assets if desired. In particular, to model cash as a perfectly liquid asset $0 \in J$ one can set $x = (x^0, \tilde{x})$ and

$$S_t(x, \omega) = x^0 + \tilde{S}_t(\tilde{x}, \omega),$$

where $\tilde{S}_t$ represents the cost function for the remaining assets $J \setminus \{0\}$. This corresponds to classical models of mathematical finance, where one of the assets serves as a “numeraire”. A more realistic option would be to assume different interest rates for lending and borrowing, as e.g. in Jouini and Kallal [34], and to set

$$S_t(x, \omega) = s^+(\omega)x^+ + s^-(\omega)x^- + \tilde{S}_t(\tilde{x}, \omega),$$

where x^+ and x^- denote the “units” invested in the lending and borrowing accounts, respectively. Here s^+ and s^- denote the unit prices of the accounts and they appreciate according to the different interest rates. For the above to make sense, we have to restrict the lending positions x^+ to be nonnegative and the borrowing positions x^- nonpositive. This brings us to portfolio constraints.

Portfolio constraints require that the portfolio $x_t(\omega)$ chosen at a given time t and state ω has to lie in a given set $D_t(\omega)$. We will assume that $\omega \mapsto D_t(\omega)$ is $\mathcal{F}_t$-measurable closed convex-valued and that $0 \in D_t$ almost surely. The measurability condition simply means that the set D_t of feasible portfolios is known to us at time t when we choose x_t. The condition $0 \in D_t$ means that we can always choose not to hold any of the traded assets.

3.1 Asset-liability management

When cash cannot be transferred quite freely in time, it is important to distinguish between payments that occur at different dates. We will denote the space of $(\mathcal{F}_t)_{t=0}^T$-adapted sequences of cash-flows by $\mathcal{M} := \{(c_t)_{t=0}^T \,|\, c_t \in L^0(\Omega, \mathcal{F}_t, P)\}$. The elements of $\mathcal{M}$ are used to model cash-flows associated with financial liabilities. A typical example would be an insurance portfolio that may obligate an insurer to claim payments over long periods of time. Simpler claims such as European options correspond to processes $c \in \mathcal{M}$ with $c_t = 0$ for all but one t.

Consider an agent whose financial position is described by a sequence of cash-flows $c \in \mathcal{M}$ in the sense that the agent has to deliver a random amount

c_t of cash at time t. Allowing c to take both positive and negative values, endowments and liabilities can be modeled in a unified manner. In particular, $-c_0$ may be interpreted as an initial endowment while the subsequent payments c_t, $t = 1, \dots, T$ may be interpreted as the cash-flows associated with financial liabilities. Problem (ALM) can be generalized to the dynamic setting as follows

$$\text{minimize} \qquad \sum_{t=0}^{T} \mathcal{V}_t(S_t(\Delta x_t) + c_t) \quad \text{over} \quad x \in \mathcal{N}_D, \tag{ALM-d}$$

where $x_{-1} := 0$,

$$\mathcal{N}_D := \{x \in \mathcal{N}, \,|\, x_t \in D_t, \ x_T = 0\},$$

and $\mathcal{V}_t : L^0(\Omega, \mathcal{F}_t, P) \to \overline{\mathbb{R}}$ are monotonic, normalized, convex functions describing the "risk/disutility/regret" from the expenditure $S_t(\Delta x_t) + c_t$ at time t.

Problem (ALM-d) may be viewed as a discrete time version of the classical Merton problem of optimal consumption [47] with illiquidity effects and a "random endowment" $-c$. On the other hand, when $\mathcal{V}_t = \delta_{L^0_-}$ for $t < T$, we can write (ALM-d) as

$$\begin{array}{ll} \text{minimize} & \mathcal{V}_T(S_T(\Delta x_T) + c_T) \quad \text{over} \quad x \in \mathcal{N}_D, \\ \text{subject to} & S_t(\Delta x_t) + c_t \le 0, \quad t = 0, \dots, T-1. \end{array} \tag{5}$$

When $T = 1$, $S_t(x, \omega) = s_t(\omega) \cdot x$, $c_0 = -w$, $c_1 = c$, we recover the one-period problem (ALM). With the traditional model of liquid markets, problem (ALM-d) can be written in terms of stochastic integrals.

Example 4 (Liquid markets and stochastic integration) *Assume that there is a perfectly liquid asset (numeraire), say $0 \in J$, such that*

$$S_t(x, \omega) = s^0_t(\omega)x^0 + \tilde{S}_t(\tilde{x}, \omega),$$
$$D_t(\omega) = \mathbb{R} \times \tilde{D}_t(\omega),$$

where $x = (x^0, \tilde{x})$ and $\tilde{S}$ and $\tilde{D}$ are the cost process and the constraints for the remaining risky assets $\tilde{J} = J \setminus \{0\}$. Expressing all costs in terms of the numeraire, we may assume $s^0 \equiv 1$. We can then use the budget constraint to substitute out the numeraire from problem (5). *Indeed, defining*

$$x^0_t = x^0_{t-1} - \tilde{S}_t(\Delta \tilde{x}_t) - c_t \quad t = 0, \dots, T-1,$$

the budget constraint holds as an equality for $t = 0, \dots, T-1$ and

$$x^0_{T-1} = -\sum_{t=0}^{T-1} \tilde{S}_t(\Delta \tilde{x}_t) - \sum_{t=0}^{T-1} c_t.$$

Substituting x^0_{T-1} in the objective (and recalling that $x_T := 0$), we can write (5) *as*

$$\text{minimize} \qquad \mathcal{V}_T\left(\sum_{t=0}^{T} \tilde{S}_t(\Delta \tilde{x}_t) + \sum_{t=0}^{T} c_t\right) \quad \text{over} \quad x \in \mathcal{N}_D.$$

In the presence of a numeraire, the timing of the payments c_t is thus irrelevant. Furthermore, in the linear case $\tilde{S}_t(\tilde{x},\omega) = \tilde{s}_t(\omega)\cdot\tilde{x}$, the cumulated trading costs can be written as the stochastic integral

$$\sum_{t=0}^{T} \tilde{S}_t(\Delta\tilde{x}_t) = \sum_{t=0}^{T} \tilde{s}_t \cdot \Delta\tilde{x}_t = -\sum_{t=0}^{T-1} \tilde{x}_t \cdot \Delta\tilde{s}_{t+1}.$$

We then obtain a discrete-time version of the utility maximization problem studied e.g. in Kramkov and Schachermayer [42] where expected utility from terminal wealth as a function of the initial endowment was studied.

We will denote the optimum value of (ALM-d) by

$$\varphi(c) := \inf\,(\text{ALM-d}).$$

This defines an extended real-valued convex function on the space $\mathcal{M}$ of $(\mathcal{F}_t)_{t=0}^T$-adapted cash-flows. When $\mathcal{V}_t = \delta_{L^0_-}$, we get $\varphi = \delta_{\mathcal{C}}$, where

$$\mathcal{C} = \{c \in \mathcal{M} \mid \exists x \in \mathcal{N}_D :\ S_t(\Delta x_t) + c_t \le 0,\ t = 0,\dots,T\}$$

is the set of claim processes that one can deliver without any cost. In the classical perfectly liquid market model described in Example 4,

$$\mathcal{C} = \{c \in \mathcal{M} \mid \exists x \in \mathcal{N}_D :\ \sum_{t=0}^{T} c_t \le \sum_{t=0}^{T-1} \tilde{x}_t \cdot \Delta\tilde{s}_{t+1}\}.$$

This set has been extensively studied in the context of arbitrage and superhedging; see [25, 16, 39] and their references. The illiquid case has been studied in [57, 59, 60].

3.2 Capital requirements

Capital requirements can be set in the dynamic setting with the same principles as in the one-period model. As in Section 2.2, we are looking for the least amount of initial capital that would allow an agent to survive a financial liability at a given level of risk. Liabilities are now described by a sequence $c \in \mathcal{M}$ of cash flows. In terms of the value function φ of problem (ALM-d), the required capital can be expressed as

$$\pi_0(c) = \inf\{\alpha \mid \varphi(c - \alpha p^0) \le 0\},$$

where $p^0 = (1, 0, \dots, 0)$. The interpretation is the same as in the one-period setting: $\pi_0(c)$ gives the least amount of initial capital one would need in order to find a hedging strategy $x \in \mathcal{N}_D$ that allows for delivering c at risk no higher than the risk of doing nothing at all. It is natural to assume that $c_0 = 0$ since a nonzero value of c_0 would just add directly to the required initial capital. The convexity of φ implies that of π_0 just like in the one-period setting; see Proposition 2.

Assuming that the minimum in (ALM-d) is attained for every $c \in \mathcal{M}$ (see Theorem 6 below) and that $\mathcal{V}_t = \delta_{L^0_-}$ as in (5), the capital requirement for a claim $c \in \mathcal{M}$ is the optimum value in the convex optimization problem

$$\begin{aligned} &\text{minimize} && S_0(x_0) \quad \text{over} \quad x \in \mathcal{N}_D, \\ &\text{subject to} && S_t(\Delta x_t) + c_t \le 0, \quad t = 1, \dots, T, \\ & && \mathcal{V}_T(S_T(\Delta x_T) + c_T) \le 0. \end{aligned} \tag{6}$$

The corresponding $x \in \mathcal{N}_D$ gives an optimal hedging strategy for the liabilities $c \in \mathcal{M}$. This approach was applied in [30] to the valuation of pension liabilities where c_t consists of the yearly pension expenditure of the Finnish private sector pension system. The case with linear market model of Example 4 and $\mathcal{V}(c) = Ev(c)$ for a convex function v with $v(\alpha) = 0$ for $\alpha \in \mathbb{R}_-$ is analyzed in [25, Section 8.2].

When $\mathcal{V}_T = \delta_{L_-}$, we get the *superhedging cost*

$$\pi_0(c) = \inf\{\alpha \,|\, c - \alpha p^0 \in \mathcal{C}\}$$

which gives the least amount of initial capital needed for delivering c without any risk of loosing money; see [59]. Extensive treatments of the superhedging cost can be found in the classical perfectly liquid model in [25, 16] and in market models with proportional transaction costs in [39].

3.3 Pricing of contingent claims

It is also straightforward to extend the pricing framework of Section 2.3 to the dynamic setting. In the one-period setting, prices of contingent claims were described in units of initial capital at time $t = 0$ that an agent would accept as a compensation for delivering a claim at a future date. In practice, however, much of trading consists of exchanging sequences of cash-flows. For example, in various swap and insurance contracts premiums are paid in sequences. The payment schedule matters since, in practice, cash cannot be transferred quite freely in time.

Consider an agent whose current financial position is characterized by a sequence $\bar{c} \in \mathcal{M}$ of cash-flows. The lowest risk he can achieve by optimally trading in the market is given by the optimum value $\varphi(\bar{c})$ of (ALM-d). Such an agent would be willing to take on additional liabilities $c \in \mathcal{M}$ in exchange for another sequence $p \in \mathcal{M}$ of cash-flows only if

$$\varphi(\bar{c} + c - p) \le \varphi(\bar{c}),$$

i.e. if the trade does not worsen the best attainable risk-return profile.

In many situations, a premium process $p \in \mathcal{M}$ is given and the aim is to find the least multiple of p that would suffice as a compensation for delivering a claim $c \in \mathcal{M}$. This leads to

$$\pi(\bar{c}; c) = \inf\{\alpha \in \mathbb{R} \,|\, \varphi(\bar{c} + c - \alpha p) \le \varphi(\bar{c})\}.$$

In the context of swap contracts, this would be the least "swap rate" the agent would accept. The usual setting in financial mathematics where premiums are paid only at the beginning is obtained as a special case when $p = (1, 0, \dots, 0)$. In particular, we recover the pricing function (4) of the one-period model when $T = 1$, $\bar{c} := (-\bar{w}, \bar{c})$, $c := (0, c)$ and $p = (w, 0)$.

Again, the function $\pi(\bar{c}; \cdot)$ is convex by the convexity of φ. Consequently, π can be analyzed much as in the one-period setting in Section 2.3; see [54] for details.

4 Duality

Convex duality has long been an integral part of financial mathematics. Classical references include Harrison and Kreps [27], Harrison and Pliska [28] and Kreps [43] where the no-arbitrage principle behind the Black–Scholes formula was related to the existence of certain "price systems". In Dalang, Morton and Willinger [14] the no-arbitrage property of a linear discrete-time market model was shown to be equivalent to the existence of a probability measure equivalent to the original measure and under which market prices are martingales; see Delbaen and Schachermayer [16] for a detailed discussion of the topic. Duality is deep-rooted in various pricing formulas, where the price of a contingent claim is expressed in terms of expectations of its cash-flows under martingale measures. Convex duality arises naturally also in portfolio optimization, where a given optimization problem is related to another much like in the classical duality frameworks of convex optimization. More generally, one can dualize a whole class of portfolio optimization problems parameterized by initial wealth and/or random future endowment; see e.g. Kramkov and Schachermayer [42] and Hugonnier and Kramkov [33] and their references.

Much like classical duality frameworks of convex optimization, duality relations in mathematical finance can often be traced back to the biconjugate theorem on convex functions in dual pairs of topological vector spaces; see Rockafellar [66]. As we have seen in the previous sections, capital requirements and prices of contingent claims can be expressed in terms of the optimum value function of an appropriately parameterized portfolio optimization problem. Similarly, dual versions of the pricing formulas can be expressed in terms of the conjugate of the optimum value function. Well-known pricing formulas and martingale characterizations of the no-arbitrage property can then be derived from these expressions as special cases.

To illustrate the ideas in the framework of Section 3, we will denote the linear space of *integrable* claim processes by

$$\mathcal{M}^1 := \{(c_t)_{t=0}^T \,|\, c_t \in L^1(\Omega, \mathcal{F}_t, P)\}.$$

This space is in *separating duality* with the linear space

$$\mathcal{M}^\infty := \{(y_t)_{t=0}^T \,|\, y_t \in L^\infty(\Omega, \mathcal{F}_t, P)\}$$

of essentially bounded adapted processes y under the bilinear form

$$\langle c,y\rangle := E\sum_{t=0}^{T} c_t(\omega)y_t(\omega).$$

Given a convex function f on $\mathcal{M}^1$, its *conjugate* is defined on $\mathcal{M}^\infty$ by

$$f^*(y)=\sup_{c\in\mathcal{M}^1}\{\langle c,y\rangle - f(c)\}.$$

Being the pointwise supremum of continuous linear functions, f^* is convex and lower semicontinuous with respect to the weak topology $\sigma(\mathcal{M}^\infty,\mathcal{M}^1)$. The classical biconjugate theorem (see [66, Theorem 5]) says that if f is proper and lower semicontinuous with respect to the L^1-norm, then it has the dual representation

$$f(c)=\sup_{y\in\mathcal{M}^\infty}\{\langle c,y\rangle - f^*(y)\}. \tag{7}$$

To apply this to the functions π_0 and π defined in Sections 3.2 and 3.3, respectively, we use the following simple fact from [54].

Lemma 5 *If the value function φ is closed, then π_0 and $\pi(\bar{c};\cdot)$ are closed as soon as they are proper.*

The conjugate of π_0 can be written as

$$\begin{aligned}
\pi_0^*(y) &= \sup_c\{\langle c,y\rangle - \pi_0(c)\}\\
&= \sup_{c,\alpha}\{\langle c,y\rangle - \alpha \,|\, \varphi(c-\alpha p^0)\le 0\}\\
&= \sup_{c,\alpha}\{\langle c,y\rangle + \alpha\langle p^0,y\rangle - \alpha \,|\, \varphi(c)\le 0\}\\
&= \begin{cases}\sup_c\{\langle c,y\rangle \,|\, \varphi(c)\le 0\} & \text{if } y_0=1,\\ +\infty & \text{otherwise.}\end{cases}
\end{aligned}$$

If there is a $c\in\mathcal{M}^1$ such that $\varphi(c)<0$ (the Slater condition), then classical Lagrangian duality gives

$$\begin{aligned}
\sup_c\{\langle c,y\rangle \,|\, \varphi(c)\le 0\} &= \inf_{\alpha>0}\sup_c\{\langle c,y\rangle - \alpha\varphi(c)\}\\
&= \inf_{\alpha>0}\alpha\sup_c\{\langle c,y/\alpha\rangle - \varphi(c)\}\\
&= \inf_{\alpha>0}\alpha\varphi^*(y/\alpha)
\end{aligned}$$

and then

$$\pi_0^*(y)=\begin{cases}\inf_{\alpha>0}\alpha\varphi^*(y/\alpha) & \text{if } y_0=1,\\ +\infty & \text{otherwise.}\end{cases} \tag{8}$$

An expression for the conjugate of $\pi(\bar{c};\cdot)$ can be derived analogously. We thus arrive at the following two-step program for deriving dual expressions for π_0 and π:

(a) establish the closedness of φ,

(b) derive expressions for φ^*.

Step (a) implies through Lemma 5 the validity of the biconjugate relation for π_0 and π while step (b) together with (8) gives an expression for the conjugate.

The above steps can be completed when the functions $\mathcal{V}_t$ in the objective of problem (ALM-d) are of the integral form

$$\mathcal{V}_t(c) = Ev_t(c) = \int_\Omega v_t(c(\omega), \omega) dP(\omega)$$

for $\mathcal{F}_t$-measurable convex normal integrands $v_t : \mathbb{R} \times \Omega \to \overline{\mathbb{R}}$ such that $v_t(\omega, 0) = 0$. The optimum value function can then be expressed as

$$\varphi(c) = \inf_{x \in \mathcal{N}, d \in \mathcal{M}} Ef(x, d, c)$$

where $Ef : \mathcal{N} \times \mathcal{M} \times \mathcal{M}^1 \to \overline{\mathbb{R}}$ is the integral functional

$$Ef(x, d, c) = \int_\Omega f(x(\omega), d(\omega), c(\omega), \omega) dP(\omega).$$

associated with the convex normal integrand

$$f(x, d, c, \omega) = \begin{cases} \sum_{t=0}^T v_t(d_t, \omega) & \text{if } S_t(\Delta x_t, \omega) + c_t \leq d_t,\ x_t \in D_t(\omega)\ t = 0, \dots, T, \\ +\infty & \text{otherwise.} \end{cases}$$

This structure allows us to extend some fundamental techniques developed for the superreplication problem in mathematical finance; see [58, 62]. First of all, the convexity of f implies that of Ef, which in turn implies the convexity of φ; see [66, Theorem 1].

Given a market model (S, D), we obtain another market model (S^∞, D^∞) by defining

$$S_t^\infty(x, \omega) = \sup_{\alpha > 0} \frac{S_t(\alpha x, \omega)}{\alpha},$$
$$D_t^\infty(\omega) = \bigcap_{\alpha > 0} \alpha D_t(\omega).$$

Indeed, the required measurability properties hold by [69, Exercises 14.54 and 14.21] while the convexity and topological properties come directly from the definitions. Moreover, the functions $S_t^\infty(\cdot, \omega)$ are positively homogeneous and the sets $D_t^\infty(\omega)$ are cones. We have $(S, D) = (S^\infty, D^\infty)$ if and only if S is sublinear and D is conical. In the language of convex analysis, $S_t^\infty(\cdot, \omega)$ is the *recession function* of $S_t(\cdot, \omega)$ and $D_t^\infty(\omega)$ is the *recession cone* of $D_t(\omega)$. An early application of recession analysis to portfolio optimization can be found in Bertsekas [7].

The following is derived in [54] from a more general result of [62].

Theorem 6 *If v_t are bounded from below by an integrable function and if*

$$\{x \in \mathcal{N}_{D^\infty} \,|\, S_t^\infty(\Delta x_t) \leq 0\}$$

is a linear space, then φ is closed in the L^1-norm topology and the infimum in (ALM-d) *is attained for every $c \in \mathcal{M}^1$.*

The boundedness assumption in Theorem 6 means that there is an integrable function $m \in L^1(\Omega, \mathcal{F}, P)$ such that $v_t(c, \omega) \geq m(\omega)$ for every $c \in \mathbb{R}$, $\omega \in \Omega$ and $t = 0, \ldots, T$. The linearity condition in Theorem 6 is a direct generalization of the classical *no-arbitrage* condition in mathematical finance; see [60, Section 4] or Example 9 below. Combining Theorem 6 with Lemma 5 establishes the lower semicontinuity of π_0 and π and thus, the validity of the dual representation (7) for them.

As to the second step (b), one may use the following result from [54] which is a straightforward application of [58, Theorem 2.2].

Lemma 7 *The conjugate of the value function φ can be expressed as*

$$\varphi^*(y) = \sigma_{\mathcal{C}^1}(y) + E\sum_{t=0}^{T} v_t^*(y_t),$$

where $\mathcal{C}^1 := \{c \in \mathcal{M}^1 \,|\, \exists x \in \mathcal{N}_D : \; S_t(\Delta x_t) + c_t \leq 0, \; x_t \in D_t\}$.

Plugging the above expression for φ^* in (8) and using the sublinearity of the support function $\sigma_{\mathcal{C}^1}$, we get

$$\pi_0^*(y) = \begin{cases} \sigma_{\mathcal{C}^1}(y) + \inf_{\alpha > 0} \alpha E \sum_{t=0}^{T} v_t^*(y_t/\alpha) & \text{if } y_0 = 1, \\ +\infty & \text{otherwise.} \end{cases} \tag{9}$$

This extends the dual representation of the *superhedging* cost given in [59, Theorem 5.2] which, in turn, extends classical superhedging formulas for liquid market models. To see this, we need a concrete expression for the support function $\sigma_{\mathcal{C}^1}$.

To this end, we will assume that the cost process S is *integrable* in the sense that $S_t(x, \cdot)$ is integrable for every $x \in \mathbb{R}^J$ and $t = 0, \ldots, T$. In the linear case $S_t(x, \omega) = s_t(\omega) \cdot x$, integrability means that the components of the price vectors s_t have finite expectations. The following result from [57], is an application of the theory of normal integrands and the Fenchel–Rockafellar duality theorem. We will denote the linear space of $(\mathcal{F}_t)_{t=0}^T$-adapted $\mathbb{R}^J$-valued integrable processes by

$$\mathcal{N}^1 := \{(w_t)_{t=0}^T \,|\, x_t \in L^1(\Omega, \mathcal{F}_t, P; \mathbb{R}^J)\}.$$

Lemma 8 *If S is integrable, then*

$$\sigma_{\mathcal{C}^1}(y) = \inf_{w \in \mathcal{N}^1} \left\{ \sum_{t=0}^{T} E(y_t S_t)^*(w_t) + \sum_{t=0}^{T-1} E\sigma_{D_t}(E_t[\Delta w_{t+1}]) \right\}$$

for every $y \in \mathcal{M}^1_+$ while $\sigma_{C^1}(y) = +\infty$ for $y \notin \mathcal{M}^1_+$. Moreover, the infimum is attained for every $y \in \mathcal{M}^1_+$.

The above results can be applied in various more specific market models. In combination with the so called Kreps–Yan theorem, Theorem 6 and Lemma 8 also yield a short proof of the "fundamental theorem of asset pricing"; see [58, 60]. We end this section by an interpretation of the above results in the classical perfectly liquid market model. More examples can be found in the above references.

Example 9 (Liquid markets) *When $S_t(x,\omega) = s_t(\omega) \cdot x$ and $D_t \equiv \mathbb{R}^J$, the linearity condition in Theorem 6 means that any self-financing trading strategy $x \in \mathcal{N}_D$ satisfies $s_t \cdot \Delta x_t = 0$ almost surely for all t. This is the classical no-arbitrage condition. When $S_t(x,\omega) = s_t(\omega) \cdot x$, we get*

$$(y_t S_t)^*(w,\omega) = \begin{cases} 0 & \text{if } w = y_t(\omega)s_t(\omega), \\ +\infty & \text{otherwise} \end{cases}$$

and when $D_t(\omega) = \mathbb{R}^J$,

$$\sigma_{D_t}(w,\omega) = \begin{cases} 0 & \text{if } w = 0, \\ +\infty & \text{otherwise.} \end{cases}$$

If s is integrable, Lemma 8 then gives

$$\sigma_{C^1}(y) = \begin{cases} 0 & \text{if } ys \text{ is a martingale}, \\ +\infty & \text{otherwise.} \end{cases}$$

Moreover, if one of the assets has constant nonzero unit price (see Example 4), then every $y \in \operatorname{dom} \sigma_{C^1}$ with $y_0 = 1$ is the density of a probability measure under which the price process s is a martingale. If $v_t = \delta_{\mathbb{R}_-}$ for all t, the dual representation (9) of the capital requirement can be written as

$$\pi_0(c) = \sup_{Q \in \mathcal{P}} E^Q \sum_{t=0}^T c_t,$$

where $\mathcal{P}$ is the set of martingale measures that are absolutely continuous with respect to P. This is a well-known expression for the superhedging cost in classical perfectly liquid market models; see e.g. [16, Section 2.4], [25, Section 5.3], [39, Chapter 2] and the references there.

5 Numerical methods

Numerical computation of capital requirements and prices of contingent claims come down to numerical solution of the asset-liability management problem (ALM-d).

Indeed, the capital requirement $\pi_0(c) = \inf\{\alpha \,|\, \varphi(c - \alpha p^0) \leq 0\}$ discussed in Section 3.2 can be approximated by line search algorithms where the optimum value $\varphi(c - \alpha p^0)$ of (ALM-d) is evaluated (approximately) for varying values of α. This approach was used in [31] to determine capital requirements for pension liabilities. The same technique can be applied to the pricing function $\pi(\bar{c}; c) = \inf\{\alpha \,|\, \varphi(\bar{c} + c - \alpha p) \leq \varphi(\bar{c})\}$ discussed in Section 3.3.

This section reviews briefly some numerical techniques for convex stochastic optimization that can be applied to (ALM) and (ALM-d) when the objective can be expressed in terms of an expectation. For simplicity, we will only consider the case $\mathcal{V}_t(c) = Ev_t(c)$ although functions of the form

$$\mathcal{V}_t(c) = \inf_{\alpha \in \mathbb{R}} \{\alpha + Ev_t(c - \alpha)\}$$

could be handled as well. The latter format covers e.g. the classical mean-variance formulation and the Conditional Value at Risk; see Section 2.1.

5.1 Static models

When $\mathcal{V}(c) = Ev(c)$ problem (ALM) can be written concisely as

$$\text{minimize} \quad Ef(x) \quad \text{over} \quad x \in X, \tag{SP}$$

where $f(x, \omega) = v(c(\omega) - s_1(\omega) \cdot x, \omega)$ and $X = \{x \in D \,|\, s_0 \cdot x \leq w\}$. This is a finite-dimensional problem but its objective involves, in general, multivariate integration. In high-dimensions, (approximate) evaluations of the objective or its gradient may be computationally demanding. In such cases, convexity becomes a valuable property.

Stochastic approximation algorithms proceed by updating a candidate solution by using "randomly sampled gradients" $v \in \partial f(x, \omega)$ where ω is randomly drawn from the distribution P and "∂" denotes subdifferentiation with respect to x. In the convex case, strong results have been obtained for certain variants of the stochastic approximation algorithm; see Nemirovski et al. [48, 44]. In particular, the "robust mirror descent stochastic approximation" algorithm constructs a random point x such that the expected value of $Ef(x)$ is guaranteed to be within an ε from the true minimum. For certain class of problems of the form (SP), the required number of gradient evaluations of f is essentially dimension-independent and grows only quadratically in $1/\varepsilon$; see [36, Proposition 1.5].

Another approach is to apply deterministic optimization algorithms to a *quadrature approximation* of (SP) obtained by replacing the underlying probability measure P by a finitely-supported measure (integration quadrature) of the form

$$P^\nu = \sum_{i=1}^{\nu} p^i \delta_{\omega^i},$$

where δ_{ω^i} denotes the Dirac measure at $\omega^i \in \Omega$ and p^i are scalars. Under such a measure, the expectation becomes a finite sum so the quadrature approximation can be written as

$$\text{minimize} \quad \sum_{i=1}^{\nu} p^i f(x, \omega^i) \quad \text{over} \quad x \in X. \tag{10}$$

This approach is useful if (a) the quadrature approximation (10) is "easy" to solve and (b) solutions of (10) are good also in the original problem (SP).

When $f(\cdot, \omega)$ are convex and p^i are positive, the quadrature approximation will be convex as well so it can be treated by numerical techniques of convex optimization; see e.g. Ben-Tal and Nemirovski [4], Nesterov [50] or Nemirovski and Juditsky [36, 37]. It is worth noting that, in the context of (ALM), once we have calculated $c(\omega^i)$ and $s_1(\omega^i)$ for a given quadrature, they can be reused when evaluating the objective of (10) and its gradient for different portfolios x. Moreover, the scenariowise components can be evaluated in parallel.

The approximation properties of (10), on the other hand, depend on the properties of the functions $f(x, \cdot)$ and the integration quadrature P^ν. In the classical Monte Carlo method, quadrature points are randomly selected from the distribution P and $p^i = 1/\nu$ for every $i = 1, \ldots, \nu$. Statistical properties of the corresponding quadrature approximations have been studied e.g. in Shapiro [72] and Shapiro and Nemirovski [73]. In quasi-Monte Carlo methods, the quadrature points are constructed by more involved techniques that achieve guaranteed accuracy for certain classes of integrands; see Novak and Woźniakowski [51, 52] and their references. Combining such techniques with efficient algorithms for convex optimization it is possible to verify the "tractability" of certain classes of stochastic optimization problems.

5.2 Tractability in the worst case setting

The *information based complexity* of a given class of stochastic optimization problems of the form (SP) is bounded by a number l if for every problem in the class, an ε-optimal solution can constructed from the information contained in f and its gradients evaluated at l points of $X \times \Omega$. Our aim is to identify classes of (SP) whose information based complexity is bounded by a polynomial of the dimensions of X and Ω and the reciprocal $1/\varepsilon$ of the required accuracy. We refer the reader to [49, 75, 51] for general treatments of problem complexity in different settings.

Stochastic approximation algorithms are examples of random algorithms that produce random solutions. The results of [48, 36] show that certain classes of (SP) are *tractable* in the *randomized setting*, i.e. the information based complexity of (SP) is bounded by a polynomial of the problem dimension and the reciprocal $1/\varepsilon$ of the required accuracy provided the quality of a random solution x is measured by the *expectation* of $Ef(x)$. The purpose of this section is to show that certain classes of (SP) are tractable also in the *worst case setting*, i.e. the information based complexity of constructing deterministic ε-optimal solutions x is bounded by a polynomial of the problem dimensions and $1/\varepsilon$. As could be expected, such "worst case" estimates require more regularity from the considered problem class.

Our strategy is to combine existing results on the complexity of convex optimization and numerical integration. The information based complexity of a class of optimization problems of the form

$$\text{minimize} \quad F(x) \quad \text{over} \quad x \in X. \tag{11}$$

is bounded by a number k if for every problem in the class, an ε-optimal solution can constructed from the information obtained by evaluating F and its gradient at k points of X. Given a seminorm L on the space of convex functions on X, we will say that a problem of the form (11) belongs to class $\operatorname{opt}(L, X)$ if it is convex and $L(F) \leq 1$. There are many well-known results on information based complexity of $\operatorname{opt}(L, X)$ when L is the Lipschitz modulus of F or of its gradient; see [36] for a recent review. Note that, as long the Lipschitz modulus of F is finite, we can scale the variables of a problem to achieve $L(F) \leq 1$. Such a scaling, of course, changes the feasible set X, which, in general, affects the complexity.

The information based complexity of a class of problems of the form

$$\text{evaluate} \quad \int_\Omega \varphi(\omega) dP(\omega) \tag{12}$$

is bounded by a number ν if every integral in the class can be evaluated to accuracy ε using information obtained by evaluating φ at ν points of Ω. Given a seminorm V on the space of integrable functions on Ω, we will say that a problem of the form (12) belongs to class $\operatorname{int}(V, P)$ if $V(\varphi) \leq 1$. There are many well-known results on information based complexity of $\operatorname{int}(V, P)$ when V is measure of "variation" or the norm on a reproducing kernel Hilbert space; see [52]. As long as $V(\varphi)$ is finite, we can scale the function φ by a constant to achieve $L(F) \leq 1$. Such a scaling, of course, affects the interpretation of the accuracy requirement ε.

It is known that optimal complexity of integration is attained by quadratures of the form

$$\sum_{i=1}^{\nu} p^i \varphi(\omega^i)$$

where p^i are real numbers; see [51, Theorem 4.7] and [52, Section 9.4]. In order to ensure that a quadrature approximation of stochastic optimization problem (SP) belongs to $\operatorname{opt}(L, X)$, we will assume that the quadrature weights p^i are nonnegative and add up to one, or in other words, that the discretized measure P^ν is a probability measure. Although most quadratures do have positive weights, there are situations where negative weights are needed for optimal complexity; see [52, Section 10.6].

We will say that the stochastic optimization problem of the form (SP) belongs to class $\operatorname{sp}(L, X, V, P)$ if

$$\sup_{\omega \in \Omega} L(f(\cdot, \omega)) \leq 1$$

and

$$\sup_{x \in X} V(f(x, \cdot)) \leq 1.$$

Note that, as long as the above suprema are finite, we can scale the variables and the objective to transform the problem into $\mathrm{sp}(L, X, V, P)$.

Lemma 10 *Let L be a seminorm on convex functions on X and let V be a seminorm on integrable functions on Ω. If*

1. *the complexity of* $\mathrm{opt}(L, X)$ *is less than $k(n, \varepsilon)$,*
2. *the complexity of* $\mathrm{int}(V, P)$ *is less than $\nu(d, \varepsilon)$ and it is attained by a probability measure P^ν,*

then the complexity of $\mathrm{sp}(L, X, V, P)$ *is less than*

$$l(n, d, \varepsilon) = \inf_{\varepsilon_1, \varepsilon_2} \{k(n, \varepsilon_1)\nu(d, \varepsilon_2) \,|\, \varepsilon_1 + 2\varepsilon_2 \leq \varepsilon\}.$$

Proof. Assume that (SP) belongs to class $\mathrm{sp}(L, X, V, P)$ and let $\varepsilon_1 > 0$ and $\varepsilon_2 > 0$ be such that $\varepsilon_1 + 2\varepsilon_2 \leq \varepsilon$. By the second assumption, there is a finitely supported probability measure P^ν with $\nu \leq \nu(d, \varepsilon_2)$ and $|E^{P^\nu}\varphi - E^P\varphi| \leq \varepsilon_2$ for all φ such that $V(\varphi) \leq 1$. Since L is sublinear, we get

$$L(E^{P^\nu} f) \leq \sum_{i=1}^{\nu} p^i L(f(\cdot, \omega^i)) \leq \sup_{\omega \in \Omega} L(f(\cdot, \omega)) \leq 1.$$

Thus, the quadrature approximation (10) belongs to $\mathrm{opt}(L, X)$ so, by the first assumption, it can be solved to accuracy ε_1 with $k(n, \varepsilon_1)$ gradient evaluations of $E^{P^\nu} f$, or equivalently, with $k(n, \varepsilon_2)\nu$ gradient evaluations of f. It thus suffices to show that the corresponding solution is an ε-solution of (SP). Indeed, the result then follows by minimizing over $\varepsilon_1 > 0$ and $\varepsilon_2 > 0$.

Let x be an ε_1-solution of (10). Since $V(f(x', \cdot)) \leq 1$ for every $x' \in X$, we have, by the choice of P^ν,

$$|E^{P^\nu} f(x) - E^P f(x)| \leq \varepsilon_2$$

and thus

$$\begin{aligned} E^P f(x) - \inf_{x' \in X} E^P f(x') &\leq [E^{P^\nu} f(x) + \varepsilon_2] - \inf_{x' \in X} [E^{P^\nu} f(x') - \varepsilon_2] \\ &= E^{P^\nu} f(x) - \inf_{x' \in X} E^{P^\nu} f(x') + 2\varepsilon_2 \\ &\leq \varepsilon_1 + 2\varepsilon_2, \end{aligned}$$

which completes the proof. □

A problem class is said to be *tractable* if its complexity is bounded from above by a polynomial of the dimension of the underlying space and the reciprocal $1/\varepsilon$ of the required accuracy. Many tractability results exist for convex optimization and for integration; see e.g. [49, 50, 36, 37] for convex optimization and [74, 51, 52] for numerical integration. Majority of the tractability results for numerical integration are attained by integration quadratures with positive weights that sum up to one.

The following corollary describes a tractable class $\mathrm{sp}(L, X, V, P)$. We denote the dimension of the decision variable x by $n = |J|$ and we assume that the domain Ω of integration is a subset of $\mathbb{R}^d$.

Corollary 11 *Assume that* $\mathrm{opt}(L, X)$ *and* $\mathrm{int}(V, P)$ *are tractable with*

$$d(n, \varepsilon) = c_1 n^{a_1}/\varepsilon^{b_1} \quad \text{and} \quad \nu(d, \varepsilon) = c_2 d^{a_2}/\varepsilon^{b_2},$$

respectively. Then $\mathrm{sp}(L, X, V, P)$ *is tractable with*

$$l(n, d, \varepsilon) = c_3 \frac{n^{a_1} d^{a_2}}{\varepsilon^{b_1+b_2}},$$

where

$$c_3 = c_1 c_2 2^{b_2} \frac{(b_1 + b_2)^{b_1+b_2}}{b_1^{b_1} b_2^{b_2}}.$$

This is achieved by a quadrature with accuracy

$$\varepsilon_2 = \frac{b_2}{b_1 + b_2} \frac{\varepsilon}{2}$$

and by solving the quadrature approximation to accuracy

$$\varepsilon_1 = \frac{b_1}{b_1 + b_2} \varepsilon.$$

Proof. By Lemma 10,

$$\begin{aligned} l(n, d, \varepsilon) &= c_1 n^{a_1} c_2 d^{a_2} \inf_{\varepsilon_1, \varepsilon_2} \{\varepsilon_1^{-b_1} \varepsilon_2^{-b_2} \,|\, \varepsilon_1 + 2\varepsilon_2 \le \varepsilon\} \\ &= c_1 n^{a_1} c_2 d^{a_2} \exp \inf_{\varepsilon_1, \varepsilon_2} \{-b_1 \ln \varepsilon_1 - b_2 \ln \varepsilon_2 \,|\, \varepsilon_1 + 2\varepsilon_2 \le \varepsilon\} \\ &= c_1 n^{a_1} c_2 d^{a_2} \exp \inf_{\varepsilon_2} \{-b_1 \ln(\varepsilon - 2\varepsilon_2) - b_2 \ln \varepsilon_2\}. \end{aligned}$$

The expression in the braces is convex in ε_1 and its derivative vanishes when

$$\varepsilon_2 = \frac{b_1}{b_1 + b_2} \frac{\varepsilon}{2}.$$

This thus gives the minimum value and the corresponding values for $\varepsilon_1 = \varepsilon - 2\varepsilon_2$ and $l(n, d, \varepsilon)$ are found by substitution. $\square$

The above shows that there do exist nontrivial classes of stochastic optimization problems that are tractable in the worst case sense. Although simple, to our knowledge these results are first of their kind. Concrete applications would require a more careful analysis of the problem structure in terms of the regularity measures L and V and complexities of the corresponding problem classes $\mathrm{opt}(L, X)$ and $\mathrm{int}(V, P)$.

5.3 Dynamic models

When $\mathcal{V}_t(c) = Ev_t(c)$, the dynamic asset-liability management problem (5) can be written as

$$\begin{array}{lll} \text{minimize} & Ev_T(S_T(\Delta x_T) + c_T) \quad \text{over} \quad x \in \mathcal{N}_D, \\ \text{subject to} & S_t(\Delta x_t) + c_t \leq 0, \quad t = 0, \dots, T-1. \end{array} \tag{13}$$

A straightforward extension of quadrature approximations to the dynamic discrete-time models leads to *scenario trees*, which may be viewed as nested quadrature approximations. Scenario trees have a long history in the field of stochastic programming but their approximation properties remain questionable. Scenario trees correspond to the "product rule" in numerical integration which quickly becomes useless when the dimensions (number of periods) increase. Asymptotic consistency properties for sequences of scenario tree approximations have been established in [53] and [55, 56], where convexity played an important role.

Galerkin methods provide a simpler computational approach which sometimes produces quite reasonable results on dynamic portfolio optimization problems. Galerkin methods are a general class of techniques for approximating infinite-dimensional optimization problems by finite-dimensional ones. The idea is to seek the best solution from a finite-dimensional subset of the original infinite-dimensional feasible set. The same idea is behind e.g. the *finite element method* which has been widely applied in physics and engineering. In the context of problem (13), the Galerkin method seeks an optimal solution among convex combinations of a finite set $\{x^i\}_{i \in I} \subset \mathcal{N}_D$ of feasible solutions (basis strategies) of (13). Such a problem can be written as

$$\text{minimize} \qquad Ev_T(S_T(\Delta \sum_{i \in I} \alpha^i x^i_T) + c_T) \quad \text{over} \quad \alpha \in X, \tag{14}$$

where

$$X = \{\alpha \in \mathbb{R}^I_+ \mid \sum_{i \in I} \alpha^i = 1\}.$$

By convexity, any convex combination of feasible solutions of (13) will automatically be feasible. Problem (14) is a *static* stochastic optimization problem and if v_t are convex functions, then (14) is convex. We can then apply the techniques outlined in the previous sections for static problems. Galerkin methods were proposed for dynamic stochastic programs in Koivu and Pennanen [41]. The underlying idea is closely related to the *affinely adjustable robust counterpart* problem proposed in Ben-Tal, Goryashko, Guslitzer and Nemirovski [3] for dynamic robust optimization.

The success of the Galerkin method rests on the properties of the basis strategies $\{x^i\}_{i \in I}$. Unlike in physics and engineering, where the dimension of the integration variable ω is often moderate, we do not have systematic techniques for generating basis strategies so that the Galerkin method would work well for arbitrary instances of (13). In practice, basis strategies are sometimes suggested by solutions of simpler problems that can be solved by dynamic programming techniques. Good results have been obtained in [41, 31] for asset liability management problems by using e.g. variants of so called "delta-hedging" and "portfolio insurance" strategies with varying parameters.

Despite its simplicity, the Galerkin method can significantly improve on the objective value produced by the basis strategies. An important advantage of the Galerkin method over e.g. the scenario tree approach is that it always produces feasible solutions that are easy to evaluate by simulation.

The loss of optimality with respect to the original problem, can be estimated if one is able to construct lower bounds for the optimum value of the original problem. In some situations, this can be done by applying Galerkin methods to a dual problem of (13). Such a technique was proposed for optimal stopping problems by Rogers [70] and Haugh and Kogan [29]. Extensions to more general problem classes are described in Pennanen [58, Remark 3.1]. Convexity is essential in guaranteeing that the optimum value of the dual can indeed approximate the optimum value of the original problem.

6 Conclusions

Many problems in financial risk management can be formulated in terms of convex optimization. Techniques of convex analysis allow for significant extensions of some fundamental results of financial mathematics to nonlinear market models with illiquidity effects and portfolio constraints. Moreover, computational techniques of convex optimization provide new possibilities in financial risk management beyond the scope of stochastic analysis alone. Although this article deals mainly with asset-liability management in financial markets, convex analysis is important also in the study of risk preferences and risk measurement; see e.g. [1, 25, 63].

Not all financial problems are convex, however. For example, fixed costs in trade execution as well as price impacts where trades affect the costs of subsequent trades lead to nonconvexities. In such situations, basic results of mathematical finance such as the "fundamental theorem of asset pricing" and martingale representations of prices of contingent claims break down. Nonconvexities arise also from the widespread use of the Value at Risk measure which lacks convexity in general. Nonconvex risk preferences are discussed in detail in [21].

Most of the results and techniques discussed in this article are far from complete. There is plenty of room for improvement and new developments. This and the recent progress in mathematical finance and industrial applications suggest that convex optimization will have a lot more to offer to financial risk management.

References

[1] Ph. Artzner, F. Delbaen, J.M. Eber, and D. Heath. Coherent measures of risk. *Mathematical Finance*, 9(3):203–228, 1999.

[2] Ph. Artzner, F. Delbaen, and P. Koch-Medona. Risk measures and efficient use of capital. *Astin Bulletin*, 39(1):101–116, 2009.

[3] A. Ben-Tal, A. Goryashko, E. Guslitzer, and A. Nemirovski. Adjustable robust solutions of uncertain linear programs. *Math. Program.*, 99(2, Ser. A):351–376, 2004.

[4] A. Ben-Tal and A. Nemirovski. *Lectures on modern convex optimization*. MPS/SIAM Series on Optimization. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2001. Analysis, algorithms, and engineering applications.

[5] A. Ben-Tal and M. Teboulle. Expected utility, penalty functions, and duality in stochastic nonlinear programming. *Management Sci.*, 32(11):1445–1466, 1986.

[6] A. Ben-Tal and M. Teboulle. An old-new concept of convex risk measures: the optimized certainty equivalent. *Math. Finance*, 17(3):449–476, 2007.

[7] D. P. Bertsekas. Necessary and sufficient conditions for existence of an optimal portfolio. *Journal of Economic Theory*, 8(2):235–247, 1974.

[8] F. Black and M. Scholes. The pricing of options and corporate liabilities. *Journal of Political Economy*, 81(3):637–654, 1973.

[9] H. Bühlmann. *Mathematical methods in risk theory*. Die Grundlehren der mathematischen Wissenschaften, Band 172. Springer-Verlag, New York, 1970.

[10] R. Carmona, editor. *Indifference pricing: theory and applications*. Princeton series in financial engineering. Princeton University Press, Princeton, NJ, 2009.

[11] U. Çetin and L. C. G. Rogers. Modelling liquidity effects in discrete time. *Mathematical Finance*, 17(1):15–29, 2007.

[12] J. Cvitanić and I. Karatzas. Convex duality in constrained portfolio optimization. *Ann. Appl. Probab.*, 2(4):767–818, 1992.

[13] J. Cvitanić and I. Karatzas. Hedging and portfolio optimization under transaction costs: a martingale approach. *Math. Finance*, 6(2):133–165, 1996.

[14] R. C. Dalang, A. Morton, and W. Willinger. Equivalent martingale measures and no-arbitrage in stochastic securities market models. *Stochastics Stochastics Rep.*, 29(2):185–201, 1990.

[15] M. H. A. Davis and A. R. Norman. Portfolio selection with transaction costs. *Math. Oper. Res.*, 15(4):676–713, 1990.

[16] F. Delbaen and W. Schachermayer. *The Mathematics of Arbitrage*. Springer Finance. Springer-Verlag, Berlin Heidelberg, 2006.

[17] M. A. H. Dempster, I. V. Evstigneev, and M. I. Taksar. Asset pricing and hedging in financial markets with transaction costs: An approach based on the Von Neumann–Gale model. *Annals of Finance*, 2(4):327–355, 2006.

[18] J. C. Dermody and E. Z. Prisman. No arbitrage and valuation in markets with realistic transaction costs. *Journal of Financial & Quantitative Analysis*, 28(1):65–80, 1993.

[19] J. C. Dermody and R. T. Rockafellar. Cash stream valuation in the face of transaction costs and taxes. *Math. Finance*, 1(1):31–54, 1991.

[20] J. C. Dermody and R. T. Rockafellar. Tax basis and nonlinearity in cash stream valuation. *Math. Finance*, 5(2):97–119, 1995.

[21] S. Drapeau and M. Kupper. Risk preferences and their robust representation. 2010.

[22] N. El Karoui and C. Ravanelli. Cash subadditive risk measures and interest rate ambiguity. *Mathematical Finance*, 19(4):561–590, 2009.

[23] I. V. Evstigneev, K. Schürger, and M. I. Taksar. On the fundamental theorem of asset pricing: random constraints and bang-bang no-arbitrage criteria. *Math. Finance*, 14(2):201–221, 2004.

[24] H. Föllmer and A. Schied. Convex measures of risk and trading constraints. *Finance Stoch.*, 6(4):429–447, 2002.

[25] Hans Föllmer and Alexander Schied. *Stochastic finance*. Walter de Gruyter & Co., Berlin, extended edition, 2011. An introduction in discrete time.

[26] L. Harris. *Trading and Exchanges, Market Microstructure for Practitioners*. Oxford University Press, 2002.

[27] J. M. Harrison and D. M. Kreps. Martingales and arbitrage in multiperiod securities markets. *J. Econom. Theory*, 20(3):381–408, 1979.

[28] J. M. Harrison and S. R. Pliska. Martingales and stochastic integrals in the theory of continuous trading. *Stochastic Process. Appl.*, 11(3):215–260, 1981.

[29] M. B. Haugh and L. Kogan. Pricing american options: a duality approach. *Oper. Res.*, 52(2):258–270, 2004.

[30] P. Hilli, M. Koivu, and T. Pennanen. Cash-flow based valuation of pension liabilities. *European Actuarial Journal*, 1:329–343, 2011.

[31] P. Hilli, M. Koivu, and T. Pennanen. Optimal construction of a fund of funds. *European Actuarial Journal*, 1:345–359, 2011.

[32] S. D. Hodges and A. Neuberger. Optimal replication of contingent claims under transaction costs. *Reviev of Futures Markets*, 8:222–239, 1989.

[33] J. Hugonnier and Kramkov. Optimal investment with random endowments in incomplete markets. *The Annals of Applied Probability*, 14(2):845–864, 2004.

[34] E. Jouini and H. Kallal. Arbitrage in securities markets with short-sales constraints. *Math. Finance*, 5(3):197–232, 1995.

[35] E. Jouini and H. Kallal. Martingales and arbitrage in securities markets with transaction costs. *J. Econom. Theory*, 66(1):178–197, 1995.

[36] A. Juditsky and A. Nemirovski. First order methods for nonsmooth convex large-scale optimization, i: General purpose methods. In *Optimization for Machine Learning*, pages 121–148. MIT Press, 2011.

[37] A. Juditsky and A. Nemirovski. First order methods for nonsmooth convex large-scale optimization, ii: Utilizing problems structure. In *Optimization for Machine Learning*, pages 149–183. MIT Press, 2011.

[38] Y. M. Kabanov. Hedging and liquidation under transaction costs in currency markets. *Finance and Stochastics*, 3(2):237–248, 1999.

[39] Y. M. Kabanov and M. Safarian. *Markets with transaction costs*. Springer Finance. Springer-Verlag, Berlin, 2009. Mathematical theory.

[40] A. J. King. Duality and martingales: a stochastic programming perspective on contingent claims. *Math. Program.*, 91(3, Ser. B):543–562, 2002. ISMP 2000, Part 1 (Atlanta, GA).

[41] M. Koivu and T. Pennanen. Galerkin methods in dynamic stochastic programming. *Optimization*, 59(3):339 – 354, 2010.

[42] D. Kramkov and W. Schachermayer. The condition on the asymptotic elasticity of utility functions and optimal investment in incomplete markets. *Annals of Applied Probability*, 9(3):904–950, 1999.

[43] D. M. Kreps. Arbitrage and equilibrium in economies with infinitely many commodities. *J. Math. Econom.*, 8(1):15–35, 1981.

[44] G. Lan, A. Nemirovski, and A. Shapiro. Validation analysis of mirror descent stochastic approximation method. *Mathematical Programming*, pages 1–34.

[45] P. Malo and T. Pennanen. Reduced form modeling of limit order markets. *Quantitative Finance*, to appear.

[46] H. M. Markowitz. *Portfolio selection: Efficient diversification of investments*. Cowles Foundation for Research in Economics at Yale University, Monograph 16. John Wiley & Sons Inc., New York, 1959.

[47] R. C. Merton. Lifetime portfolio selection under uncertainty: The continuous-time case. *Review of Economics and Statistics*, 3:373–413, 1969.

[48] A. Nemirovski, A. Juditsky, G. Lan, and A. Shapiro. Robust stochastic approximation approach to stochastic programming. *SIAM J. Optim.*, 19(4):1574–1609, 2008.

[49] A. S. Nemirovsky and D. B. Yudin. *Problem complexity and method efficiency in optimization*. A Wiley-Interscience Publication. John Wiley & Sons Inc., New York, 1983. Translated from the Russian and with a preface by E. R. Dawson, Wiley-Interscience Series in Discrete Mathematics.

[50] Yu. Nesterov. *Introductory lectures on convex optimization*, volume 87 of *Applied Optimization*. Kluwer Academic Publishers, Boston, MA, 2004. A basic course.

[51] E. Novak and H. Woźniakowski. *Tractability of multivariate problems. Vol. 1: Linear information*, volume 6 of *EMS Tracts in Mathematics*. European Mathematical Society (EMS), Zürich, 2008.

[52] E. Novak and H. Woźniakowski. *Tractability of multivariate problems. Volume II: Standard information for functionals*, volume 12 of *EMS Tracts in Mathematics*. European Mathematical Society (EMS), Zürich, 2010.

[53] P. Olsen. Discretizations of multistage stochastic programming problems. *Math. Programming Stud.*, (6):111–124, 1976. Stochastic systems: modeling, identification and optimization, II (Proc. Sympos., Univ. Kentucky, Lexington, Ky.,1975).

[54] T. Pennanen. Indifference pricing in illiquid markets. *manuscript*.

[55] T. Pennanen. Epi-convergent discretizations of multistage stochastic programs. *Math. Oper. Res.*, 30(1):245–256, 2005.

[56] T. Pennanen. Epi-convergent discretizations of multistage stochastic programs via integration quadratures. *Math. Programming, Series B*, 116(1):461–479, 2009.

[57] T. Pennanen. Arbitrage and deflators in illiquid markets. *Finance and Stochastics*, 15(1):57–83, 2011.

[58] T. Pennanen. Convex duality in stochastic optimization and mathematical finance. *Mathematics of Operations Research*, 36(2):340–362, 2011.

[59] T. Pennanen. Superhedging in illiquid markets. *Mathematical Finance*, 21(3):519–540, 2011.

[60] T. Pennanen. Dual representation of superhedging costs in illiquid markets. *Mathematics and Financial Economics*, to appear.

[61] T. Pennanen and I. Penner. Hedging of claims with physical delivery under convex transaction costs. *SIAM Journal on Financial Mathematics*, 1:158–178, 2010.

[62] T. Pennanen and A.-P. Perkkiö. Stochastic programs without duality gaps. *Mathematical Programming*, to appear.

[63] G. Ch. Pflug and W. Römisch. *Modeling, Measuring and Managing Risk*. World Scientific, Singapore, 2007.

[64] S. R. Pliska. *Introduction to Mathematical Finance: Discrete Time Models*. Blackwell, Malden, MA, 1997.

[65] R. T. Rockafellar. *Convex analysis*. Princeton Mathematical Series, No. 28. Princeton University Press, Princeton, N.J., 1970.

[66] R. T. Rockafellar. *Conjugate duality and optimization*. Society for Industrial and Applied Mathematics, Philadelphia, Pa., 1974.

[67] R. T. Rockafellar and S.P. Uryasev. Optimization of Conditional Value-at-Risk. *Journal of Risk*, 2:21–42, 2000.

[68] R. T. Rockafellar and S.P. Uryasev. The fundamental risk quadrangle in risk management, optimization and statistical estimation. *working paper*, 2011.

[69] R. T. Rockafellar and R. J.-B. Wets. *Variational analysis*, volume 317 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 1998.

[70] L. C. G. Rogers. Monte Carlo valuation of American options. *Math. Finance*, 12(3):271–286, 2002.

[71] W. Schachermayer. The fundamental theorem of asset pricing under proportional transaction costs in finite discrete time. *Math. Finance*, 14(1):19–48, 2004.

[72] A. Shapiro. Monte Carlo sampling methods. In *Stochastic programming*, volume 10 of *Handbooks Oper. Res. Management Sci.*, pages 353–425. Elsevier, Amsterdam, 2003.

[73] A. Shapiro and A. Nemirovski. On complexity of stochastic programming problems. In *Continuous optimization*, volume 99 of *Appl. Optim.*, pages 111–146. Springer, New York, 2005.

[74] I. H. Sloan and H. Wozniakowski. When are quasi-monte carlo algorithms efficient for high dimensional integrals? *J. Complexity*, 14:1–33, 1997.

[75] J. F. Traub and A. G. Werschulz. *Complexity and Information*. Cambridge University Press, Cambridge, 1998.