Risk Management by Markov Decision Processes

by

You Liang

A thesis submitted to the Faculty of Graduate Studies of
The University of Manitoba
in partial fulfilment of the requirements for the degree of

Doctor of Philosophy

Department of Statistics
The University of Manitoba
Winnipeg

Abstract

In the financial world, risk is both unpredictable and unavoidable. However risk brings opportunity for profits such as in the stock market, although in most situations risk is negative and dangerous. A very important and powerful tool in the study of mathematical finance is the model of Markov decision processes which are particularly useful because they are general enough to model a variety of natural processes in practical situations involving uncertainty, and mathematically tractable to derive explicit solutions and structural properties that are useful to guide decision-making in such situations.

My PhD research in the area of risk management is focused on three major problems: (1) risk sensitive control of partially observable Markov decision processes, (2) conditional and dynamic risk measures, (3) conditional and dynamic deviation measures.

The classic model of Markov decision processes assumes completely observable states of the Markov processes and is risk neutral. Our first part of the thesis is to extend the classic model of Markov decision processes simultaneously in two directions: partially observable states and risk sensitivity. The motivation is very natural. More often than not in practical situations involving uncertainty, states of the underlying processes may only be observed through some signals.

On the other hand, attitudes towards risk (such at risk-seeking, risk-averse and risk-neutral) definitely affects the choices of decision-making.

Another direction of extending the classic model of Markov decision processes is to incorporate risk measures with the optimality criteria of the model. Measuring and controlling risk is very important for financial institutions including banks and insurance companies. Our second part of the thesis is to use the model of Markov decision processes to characterize and derive new forms of dynamic risk measures.

The standard deviation may be the first deviation measure used in mathematical finance, and can be generalized in many ways. The third part of the thesis is to apply the model of Markov decision processes to characterize and derive a sequence of dynamic deviation measures. To the best of our knowledge, this is the first time that the concept of dynamic deviation measures is introduced.

Acknowledgements

First of all, I thank my thesis co-advisors Dr. Xikui Wang and Dr. Yanqing Yi for their patient guidance and constant support during this research. Without their helpful comments and motivation, this thesis would not have been possible. I also thank the members of my PhD Thesis Advisory Committee, Dr. Brad Johnson, Dr. A. Thavaneswaran and Dr. Xuemiao Hao, for their support and constructive criticism of my research, and the external examiner Dr. Yiqiang Zhao (from Carleton University) for his constructive criticism.

I gratefully acknowledge financial supports from the Faculty of Graduate Studies, the Faculty of Science, the Department of Statistics, and from Dr. Wang's research grants and Dr. Yi's research grants.

Finally I thank all members of the Department of Statistics and many friends for their support and encouragement.

To My Parents

Contents

Chapter 1

Introduction

Risk exists everyday and everywhere, in many different forms and with very different consequences and impacts. Risk is an uncertain event in the future. To minimize its consequence, it is important to understand risk, define and characterize it, manage it and minimize its potential impact. Risk management is an important area of research in mathematical finance. A very useful tool for risk management is the model of Markov decision processes. In this chapter, we introduce the backgrounds of mathematical finance, risk management and Markov decision processes.

1.1 Mathematical Finance

There are two dimensions of any financial investment in the face of uncertainty: mean return and risk of the returns. The two goals are to maximize the mean return and to minimize the variance of the returns. However for this multi-objective optimization problem, there is in general no uniquely defined best decision to optimize in both dimensions. We have to compromise. The set of optimal solutions which are Pareto-efficient in the sense of these two objectives is called the mean-risk efficient frontier. In some models for optimal financial decision making the two dimensions are often mixed by introducing a nondecreasing concave utility function. Risk aversion, i.e. the degree to which the risk dimension is taken into account, can be modelled by the negative curvature of the utility function.

Financial investment is inherently risky in that its future values follow an unknown distribution or a known distribution with unknown parameters. As common practice, an investor normally assesses the likely future behaviour of his investments, or portfolio, by constructing a model which gives a prediction of future values. The investment risk is then divided into a known distribution risk and an unknown model risk. The first means that observations from a known distribution are random, and the second implies that the chosen model may be incorrect. These are often referred to as risk and uncertainty respectively. Although some degree of uncertainty may be factored into the model, its unmeasurable nature can make the model itself particularly problematic. Any portfolio optimization method will be subject to some degree of uncertainty because the outcomes occur in the unpredictable future. Factors that make up the model risk may include systemic risk, regulatory risk, or any other potential causes for capital loss (Rebonato, 2001).

Clearly in the investment problem, the attitude of risk aversion of an investor dictates the preference of high investment values against low values. This may be achieved by ordering the available decisions according to some function of the values. One method is to use the utility function (Fishburn, 1970). Another method uses a risk preference which is expressed by the ratio of a measure of the expected return and a measure of risk such as the variance or standard deviation. This approach involves behavioural preferences which are context dependant (Tversky and Simonson, 1993).

Since its birth as an independent branch of social sciences, the three eras of mathematical finance (Szegö, 2004) are mean-variance models, continuous-time models and risk measures. The first revolution started when Markowitz (1952, 1959) proposed to measure risk associated with the return of each investment by maximizing the expected portfolio return and minimizing the standard deviation of the returns. Markowitz also proved that normally and log-normally distributed share returns make the return variance compatible with the utility functions. Markowitz focused on the standard deviation as a risk measure, and it is not difficult to extend his idea to other dispersion measures, such as the absolute deviation, to measure risk levels and to be compatible with the utility functions.

Another important contribution in this period was simultaneously presented by Roy (1952). He maximized the expected return of a portfolio of stocks but the risk level was measured by the probability of losing money. Since this probability is bounded from above by using variances and the Tchebycheff inequality, the variance acts as the risk measure in practice.

Sharpe (1964) invented the Capital Asset Pricing Model (CAPM) which is one of the most important equilibrium models in financial economics. The CAPM was built on the earlier work of Harry Markowitz so as to introduce the systematic risk common to all securities and unsystematic risk associated with individual assets. Once again risk levels were measured with variances. Sharpe, Markowitz and Miller jointly received the 1990 Nobel Memorial Prize in Economics for their contributions to the field of financial economics.

Ross (1976) published the Arbitrage Pricing Theory (APT), a key innovation in financial economics that partially extends the concept of CAPM and includes other economic factors (such as inflation, interest rates, exchange rates and oil prices). He was awarded the Graham & Dodd Prize for his distinguished research on a wide range of work in mathematical finance theory including the theory of agency, the binomial model of option pricing, and the Cox-Ingersoll-Ross term structure model of interest rates.

More authors have dealt with the variance as a risk measure and generalized the contributions above by considering strictly weaker assumptions about the asset return behaviour, the utility function or other involved economic properties. The Stochastic Discount Factor (SDF) is a key notion developed from the Riesz Representation Theorem, which leads to some major formulas for both CAPM and

APT. Another question recently generating a growing interest is the compatibility between a dispersion measure and the Second Order Stochastic Dominance (SOSD). As explained above, Markowitz's theory is only compatible with normally and log-normally distributed returns and the usual utilities. However, empirical studies have indicated that the presence of asymmetries, heavy tails and extreme values is becoming more and more evident. In a non-Gaussian world, the arguments of Markowitz do not work any more, and the variance is not an appropriate risk measure. Indeed, it is not compatible with the SOSD. By minimizing the standard deviation one could find a portfolio that does not maximize any rational utility function. Recent research has shown that this caveat may be overcome by using alternative dispersion measures, such as the absolute deviation and semideviation, the standard semideviation etc. See (Ogryczak and Ruszczyński, 1999, 2001, 2002) for further details.

The second revolution was started by Robert Merton, Fisher Black and Myron Scholes (see (Black and Scholes, 1973) and (Merton, 1989)), and it can be labeled as "continuous-time finance". It made it possible to crack many problems associated with option-pricing and other derivatives. The concept of contingent claim, essential in modern finance, is a byproduct of these theories. According to the Black and Scholes theory, the risk level generated by any derivative security may be neutralized by trading δ units of the underlying asset, where δ is the partial derivative (sensitivity) of the derivative security price with respect to the underlying asset price. Thus, once again risk levels can be measured by sensitivities. Other famous "sensitivities" are Γ (sensitivity of δ with respect to the asset price), θ (sensitivity of the portfolio price with respect to time), ρ (sensitivity of the portfolio price with respect to the riskless interest rate) and v (sensitivity of the portfolio price with respect to the asset volatility). Some more "sensitivities" have been introduced recently, though the basic ideas remain the same as above.

The third revolution is much more recent and started in 1997. It was motivated by the growing development of the financial and actuarial sectors, which provoked a new point of view when measuring risk levels. The new approach must consider the necessities of regulators, who must provide rules guaranteeing the stability of the system, supervisors, who must control the industrial activities respect the legal framework, and public or private companies, that must manage the wealth of their customers. In the European Union the set of rules that the industry must respect are mainly contained in Basle II (finance) and Solvency II (insurance). They provide the method that any corporation must follow when computing its "capital reserves", i.e., additional capital that will be devoted to overcome those periods characterized by losses of the economic activity. The size of the appropriate reserve may be considered as the risk level associated with the firm (or its activity). From a mathematical perspective this reserve will be understood as a real-valued function on an $L^k(\Omega, \mathcal{F}, P)$ space, where $k \in [1, \infty]$, and $(\Omega, \mathcal{F}, P)$ is an arbitrary probability space. The study of these functions, that we will call modern risk measures, is the focus of the Modern Risk Management.

Indeed in the last few years, there has been a great momentum in research on this subject, which has touched nine different, but interconnected, problems: critique of current risk measures, definition of risk measure, construction of (coherent) risk measures, rationality of insurance premium, conditional CAPM (Franke et al., 2000), “good deals”, asset pricing in incomplete markets, generalized hyperbolic Lévy processes (Geman, 2002), copulas for the study of co-dependence and crash-prediction methods (Sornette and Johansen, 2001).

1.2 (Partially Observable) Markov Decision Processes

Since the pioneering works of Bellman (1957) and Howard (1960) in the early 1950’s, Markov decision processes (MDPs) have becoming the most important research area in operations research. MDPs are the marriage between Markov processes and dynamic programming, combining the ideas of stochastic dynamic modelling and optimal sequential control. Such a happy marriage makes the MDPs not only mathematically tractable but also widely applicable, and hence most powerful and popular. Alternatively, they are known as controlled Markov processes, stochastic dynamic programming, Markov decision programming, and Markov control processes. The MDP models capture the two fundamental features underlying risk management: a practical situation involving uncertainty (which is characterized by the risk process) and the desire for its control.

Consider a discrete time Markov decision process consisting of five elements: the state space S, the feasible action space A, the state transition law $P(j|i,a), a \in A, i, j \in S$, the one step random payoff function $r(i,a), i \in S, a \in A$ and an optimality criterion. At any time point $n = 1, 2, \cdots$, the state $s_n \in S$ of the Markov process summarizes all necessary information required for making a decision at this state. If action $a_n \in A$ is taken at state s_n, then two things happen: a random payoff $r(s_n, a_n)$ is collected and the controlled Markov process moves to state $s_{n+1} \in S$ at the next time point with probability $P(s_{n+1}|s_n, a_n)$. A history-dependent, randomized strategy is a sequence of decision rules $\pi = (\pi_1, \pi_2, \cdots)$ such that action a_n is taken with probability $\pi_n(a_n|s_n, h_n)$ for any given history $h_n = (s_1, a_1, \cdots, s_{n-1}, a_{n-1})$ of past states and actions. If $\pi_n(a_n|s_n, h_n) = \pi_n(a_n|s_n)$ does not depend on the history h_n for all n, then the strategy is Markovian. If $\pi_n(a_n|s_n, h_n) = 1$ for a particular action a_n for all n (that is, π_n is degenerate for all n), then the strategy is called deterministic. A Markovian, deterministic strategy is stationary if it does not depend on time.

Given a distribution p_1 (which may be degenerate) of the initial state s_1 of the controlled Markov process and the state transition probabilities $P(s_{n+1}|s_n, a_n)$, a strategy π defines a probability measure $P^{\pi}_{p_1}$ on the random sample path $(s_1, a_1, s_2, a_2, \cdots)$, which in turn defines an expectation $\mathrm{E}^{\pi}_{p_1}$. The optimality criterion is a function of the initial distribution p_1 and the strategy π, and is defined by the expected value (under the expectation $\mathrm{E}^{\pi}_{p_1}$) of a function of the sequence $r(s_n, a_n), n = 1, 2, \cdots$, of random payoffs. Commonly used criteria include (1) the finite N horizon criterion

$$V_N^\pi(p_1) = \mathrm{E}_{p_1}^\pi \left[\sum_{n=1}^{N-1} r(s_n, a_n) + R(s_N) \right]$$

where $R(s)$ is a random terminal payoff, (2) the infinite horizon β-discounted criterion

$$W_\beta^\pi(p_1) = \lim_{N\to\infty} \mathrm{E}_{p_1}^\pi \left[\sum_{n=1}^{N} \beta^{n-1} r(s_n, a_n) \right] = \mathrm{E}_{p_1}^\pi \left[\sum_{n=1}^{\infty} \beta^{n-1} r(s_n, a_n) \right]$$

if the limit exists, where $0 < \beta < 1$, and (3) the infinite horizon average criterion

$$G^\pi(p_1) = \lim_{N\to\infty} \frac{1}{N} \mathrm{E}_{p_1}^\pi \left[\sum_{n=1}^{N} r(s_n, a_n) \right]$$

if the limit exists. The goal is to find an optimal strategy such that the controlled stochastic dynamic system performs optimally with respect to a predetermined optimality criterion. This means that we maximize the criteria if the payoffs are rewards, and we minimize the criteria if the payoffs are costs.

There are usually four major research objectives of MDPs with any given criterion: (1) demonstrate the existence of an optimal strategy, the optimality of special strategies (such as stationary ones), and the conditions under which an optimal strategy exists, (2) characterize the optimal strategy and the optimal value, normally by an optimality equations (which is often derived by applying the principle of dynamic programming), (3) investigate important theoretical and structural properties of the optimal strategy, normally by means of the optimality equation, which often assist in computing the optimal strategy and the optimal value, and (4) apply MDPs to solve real world research problems. MDPs have been widely applied to many areas of research, including management, finance and economics, queueing theory, and many more. Hernández-Lerma and Lasserre (1996) and Puterman (2014) are all good reference books on Markov decision processes.

MDPs have been generalized in a number of directions, including partially observable Markov decision processes (POMDPs). In a model of POMDPs, the true state of the underlying Markov process is unobservable. Instead, a signal is observed and is named the information state, based on which a decision is made. Another generalization is to incorporate the risk attitude with the optimality criterion, in the form of a utility function. These models are named risk sensitive Markov decision processes (RSMDPs). A third generalization is to incorporate dynamic risk measures with the optimality criterion. Moreover, dynamic deviation measures may also be incorporated.

The focus on the thesis is to work on these three extensions.

1.3 Modern Risk Measures

Risk measures are financial tools for quantitatively determining the minimum capital reserves. The financial institutions (such as banks, insurance companies and others) are required to maintain such reserves in order to ensure their financial stability. The uncertainty in the future value of a portfolio is modelled by a random variable, to which a functional of a risk measure is applied. Mathematically, a risk measure can be described by a functional from a certain function space to the space of real numbers. Let $(\Omega, \mathcal{F}, P)$ be a given probability space, where the sample space Ω represents the set of states of nature. Throughout the whole thesis, all equalities and inequalities applied to random variables are in the sense of almost surely. Without loss of generality, let L^∞ be the space of essentially bounded functions on $(\Omega, \mathcal{F}, P)$, and $\mathbb{R}$ be the space of all real numbers. A risk measure is defined as a real-valued functional $\rho : L^\infty \to \mathbb{R}$. When $X \in L^\infty$ represents the financial state of a financial institution, $\rho(X)$ gives a measure of risk at this financial state.

Value at Risk (VaR) has been used as an industry standard, but is often criticized for encouraging the accumulation of shortfall risk in particular scenarios. For the random variables X defined on a probability space $(\Omega, \mathcal{F}, P)$, VaR at level $\alpha \in (0, 1)$ is defined by

$$\mathrm{VaR}_\alpha(X) = \inf\{m \in R | P(X + m < 0) \leq \alpha\}.$$

The $100(1-\alpha)\%$ VaR of a random return X is defined as the negative of the α-quantile of the return distribution F_X

$$\mathrm{VaR}_\alpha(X) = -\inf\{x | F_X(x) > \alpha\} = -F_X^{-1}(\alpha),$$

where F_X^{-1} is the inverse function of the distribution function F_X. Suppose that X follows a normal distribution $N(\mu, \sigma^2)$. The quantile function of X is $F_X^{-1}(\alpha) = \mu + \sigma\Phi^{-1}(\alpha)$, where $\Phi^{-1}(\alpha)$ is the quantile function of the standard normal distribution. Hence

$$\mathrm{VaR}_\alpha(X) = -\mu - \sigma\Phi^{-1}(\alpha).$$

However, the following example shows that VaR is not subadditive. Consider two assets X and Y that are normally distributed, but subject to the occasional independent shocks:

$$X = \epsilon + \eta, \quad \epsilon \overset{iid}{\sim} N(0,1), \eta = \begin{cases} 0 & \text{with probability } 0.991 \\ -10 & \text{with probability } 0.009 \end{cases}$$

In this case, the 99% VaR is 3.1, since the probability that η is -10 is less than 99%. Suppose that Y has the same distribution independently from X, and that we formulate an equally weighted portfolio of X and Y. In that case, the 99% portfolio VaR is 9.8, because for $X + Y$ the probability of getting the -10 draw for either X or Y is much higher. Therefore,

$$\mathrm{VaR}(X + Y) = 9.8 > \mathrm{VaR}(X) + \mathrm{VaR}(Y) = 3.1 + 3.1 = 6.2$$

To overcome this deficiency, appropriate alternatives of risk measures have been proposed that are more reasonable. Firstly, a system of desirable axioms for monetary risk measures are specified. Secondly, risk measures are characterized that satisfy these axioms. Thirdly, suitable examples are identified and the representation of risk measures in terms of their acceptance sets is investigated. In the next few subsections, we introduce the various sets of axioms, define coherent and convex risk measures, and discuss representation of monetary, coherent and convex risk measures in terms of their acceptance sets. The general dual representation for convex and coherent risk measures are discussed, together with various examples. In many practical situations, it is reasonable and desirable to assume that a risk measure is law-invariant, meaning that the risk measure depends on the randomness of the future value only through its probability law.

1.3.1 Coherent Risk Measures

Systematic research on risk measures was initiated by Artzner et al. (1997, 1999), who provided an axiomatic analysis of risk assessment in terms of capital reserves and introduced the concept of coherent risk measures. In their framework, several properties of risk measures (including monotonicity, positive homogeneity, subadditivity) are required.

Artzner et al. (1999) not only defined risks (including both market and non-market risks) but also provided and justified an axiomatic system for analyzing, constructing and implementing coherent risk measures. These risk measures can be regarded as extra capital requirements to regulate the risk assumed by market participants, traders, and insurance underwriters. They suggested that risk be defined as future values instead of changes in value between two dates, because they believed that risk is related to the variability of the future value of a position due to market changes or more generally to uncertain events. They based their future wealth approach for risk on the principle of "bygones are bygones".

The underlying function space $\mathcal{G}$ is defined as the space of real-valued random variables on the set of future states of nature Ω, interpreted as possible future values of positions or portfolios currently held. This set Ω consists of a fixed set of scenarios of future financial states. A real-valued random variable X on Ω represents a risk, which for example, can be the discounted value of the portfolio or the sum of its P&L and some economic capital. One goal of risk management is to determine a number $\rho(X)$ that quantifies the risk of the financial position X and can serve as a capital requirement. That is, $\rho(X)$ is the minimal amount of capital that, if added to the financial position X and invested in a risk-free manner, makes the financial position acceptable.

Artzner et al. (1999) discussed a first crude, but crucial, measurement of the risk of a financial position as whether or not its future value belongs or does not belong to a subset of "acceptable risks". Being acceptable means that no additional capital is required. This subset of acceptable risks may be decided by a supervisor such as a regulator who considers the unfavourable states when allowing a risky position that may draw on the resources of the government, or an exchange clearing firm which has to make good on the promises to all parties of transactions being securely completed, or an investment manager who knows that his firm has basically given to its traders an exit option in which the strike price consists in being fired in the event of big trading losses on one's position (Artzner et al., 1999).

Correspondingly, an "unacceptable risk" is a position with an unacceptable future net worth. One reasonable remedy is to look for some commonly accepted instruments (such as cash) that, when added to the current position, make its future value acceptable to the regulator or supervisor. Therefore one good candidate for a risk measure of the initially unacceptable financial position is the current cost of getting enough of these instruments.

Artzner et al. (1999) considered only the case of a finite Ω with $n = card(\Omega)$. In this case, we can identify $\mathcal{G}$ virtually as $\mathbb{R}^n$. In their framework, L_+ denotes the nonnegative elements in $\mathcal{G}$ and L_- denotes the negative elements. Furthermore, $L_{--} = \{X \in \mathcal{G} | for\ each\ \omega \in \Omega, X(\omega) < 0\}$. They also assumed a static model with two time points: 0 and a terminal time $T > 0$. The total number of currencies in the market is I. Expressed in currency $i = 1, 2, \cdots, I$, they denoted $\mathcal{A}_{i,j}, j \in J_i$, the set of final net worths in currency i that are accepted by regulator j. Then the intersection $\mathcal{A}_i = \cap_{j \in J_i} \mathcal{A}_{i,j}$, generically denoted as $\mathcal{A}$ in the following framework, is the set of all accepted positions for currency i.

The following is the collection of axioms for acceptance sets as assumed in (Artzner et al., 1999):

(AS1) The acceptance set $\mathcal{A}$ contains L_+.

(AS2) The acceptance set $\mathcal{A}$ does not intersect the set L_{--}. Or strongly,

(AS2') The acceptance set $\mathcal{A}$ satisfies $\mathcal{A} \cap L_- = \{0\}$.

(AS3) The acceptance set $\mathcal{A}$ is convex.

(AS4) The acceptance set $\mathcal{A}$ is a positively homogeneous cone.

The set of acceptable future net worths is only the primitive object to describe acceptance or rejection of a risk. Artzner et al. (1999) investigated the correspondence between the acceptance sets and the risk measures. A positive value of the risk measure $\rho(X)$ is interpreted as the minimum extra cash the agent has to add to the risky position X and invest prudently. A negative value of the risk measure $\rho(X)$ means that the cash amount $-\rho(X)$ can be withdrawn from the position or it can be received as restitution, as in the case of organized markets for financial futures. Suppose that r is a given total rate of return on a reference instrument. The risk measure associated with a acceptance set $\mathcal{A}$ is defined in (Artzner et al., 1999) as a functional $\rho_{\mathcal{A},r}$ from $\mathcal{G}$ to $\mathbb{R}$ such that $\rho_{\mathcal{A},r}(X) = \inf\{m | m \cdot r + X \in \mathcal{A}\}$. Conversely given any risk measure ρ, the acceptance set $\mathcal{A}_\rho$ associated with ρ is defined as $\mathcal{A}_\rho = \{X \in \mathcal{G} | \rho(X) \leq 0\}$.

Artzner et al. (1999) then introduced the following axiomatic system for coherent risk measures.

Translation invariance: For all real-valued random variables X and all real numbers (i.e., cashes) α, we have $\rho(X + \alpha \cdot r) = \rho(X) - \alpha$, where r is the given return rate on a reference instrument.

Subadditivity: For any two real-valued random variables X and Y, we have $\rho(X + Y) \leq \rho(X) + \rho(Y)$.

Positive homogeneity: For all real-valued random variables X and all real numbers λ, we have $\rho(\lambda X) = \lambda\rho(X)$.

Monotonicity: For any two real-valued random variables X and Y such that $X \leq Y$, we have $\rho(Y) \leq \rho(X)$.

Relevance: For all real-valued random variables $X \neq 0$ such that $X \leq 0$, we have $\rho(X) > 0$.

A mapping $\rho : \mathcal{G} \to \mathbb{R} \cup \{+\infty\}$ is called monetary risk measure if $\rho(0)$ is finite and satisfies the Axioms of monotonicity and translation invariance. Here, we assume discounted quantities. See (EI Karoui and Ravaelli, 2009) for a discussion of forward risk measures and interest rate ambiguity.

A risk measure ρ is said to be coherent if it satisfies the four axioms of translation invariance, subadditivity, positive homogeneity, and monotonicity (Artzner et al., 1999). Finally, the correspondence between the acceptance sets satisfying the axioms and the risk measures satisfying the axioms is established. Specifically, if $\mathcal{F}$ is a set satisfying Axioms AS1 to AS4, then the risk measure $\rho_{\mathcal{F},r}$ is coherent and $\mathcal{A}_{\rho_{\mathcal{F},r}} = \bar{\mathcal{F}}$, the closure of $\mathcal{F}$. Conversely, if ρ is a coherent risk measure, then the acceptance set $\mathcal{A}_\rho$ (called the acceptable set of ρ) is closed and satisfies Axioms AS1 to AS4. Moreover, $\rho = \rho_{\mathcal{A}_\rho,r}$. The acceptance set $\mathcal{F}$ satisfies the stronger Axiom AS2' if and only if its corresponding risk measure satisfies the Relevance Axiom. Furthermore, it is true that $\mathcal{A}_\rho \cap \mathbb{R} \neq \emptyset$, $\inf\{m \in \mathbb{R} | X + m \in \mathcal{A}_\rho\} > -\infty$ for all $X \in \mathcal{G}$, and that $X \in \mathcal{A}_\rho$, $Y \in \mathcal{G}$ and

$X \leq Y$ all together imply that $Y \in \mathcal{A}_\rho$.

For example, the Conditional Value at Risk (CVaR), which is also called Average Value at Risk (AVaR), Expected Shortfall (ES), or Tail Value at Risk (TVaR), at level $\lambda \in (0,1]$, defined as $\mathrm{CVaR}_\lambda(X) = \frac{1}{\lambda}\int_0^\lambda \mathrm{VaR}_\alpha(X)d\alpha$, is a coherent risk measure. If we assume that the return distribution F_X is continuous at $-\mathrm{VaR}_\lambda(X)$, the CVaR equals to the following conditional expectation

$$\mathrm{CVaR}_\lambda(X) = -\mathrm{E}[X|X \leq -\mathrm{VaR}_\lambda(X)]. \tag{1.1}$$

This requires a more subtle definition to be able to handle possible discontinuities. The CVaR can be represented through a minimization formula as

$$\mathrm{CVaR}_\lambda(X) = \min_{C\in\mathbb{R}} \left\{ C + \frac{1}{\lambda}\mathrm{E}[X+C]^- \right\} \tag{1.2}$$

$$= \min_{C\in\mathbb{R}} \left\{ C + \frac{1}{\lambda}\mathrm{E}[-X-C]^+ \right\} \tag{1.3}$$

where $(x)^+ = \max\{x, 0\}$ and $(x)^- = \max\{-x, 0\}$. It turns out that this formula has an important application in the optimal portfolio problems using CVaR as the risk measure. The above definition of CVaR as a minimization formula was first studied in (Pflug, 2000), and then further investigated in (Rockafellar and Uryasev, 2002). Their approach allows the treatment of both continuous and discontinuous F_X. Rockafellar and Uryasev (2002) proved that (1.2) is equivalent to the conditional expectation method (1.1), by replacing the conditional expectation in (1.1) with the "generalized λ-tail expectation"

$$\mathrm{CVaR}_\lambda(X) = -\int_{-\infty}^{\infty} x dF_X^\lambda(x), \quad \text{where } F_X^\lambda(x) = \begin{cases} \lambda^{-1}F^\lambda(x) & \text{if } F^\lambda(x) < \lambda, \\ 1 & \text{if } F^\lambda(x) \geq \lambda. \end{cases}$$

A different track was followed by Acerbi (2002), by showing that CVaR can equivalently be expressed as a average VaR

$$\mathrm{CVaR}_\lambda(X) = \frac{1}{\lambda}\int_0^\lambda \mathrm{VaR}_\alpha(X)d\alpha. \tag{1.4}$$

If $F(F^{-1}(\lambda)) = \lambda$, then

$$\mathrm{CVaR}_\lambda(X) = -\frac{1}{\lambda}\int_0^\lambda F_X^{-1}(\alpha)d\alpha = -\frac{1}{\lambda}\int_{-\infty}^{F_X^{-1}(\lambda)} x dF_X(x).$$

The CVaR has not only a good practical meaning, but also sound mathematical properties. CVaR satisfies the coherency axioms introduced in (Artzner et al., 1999). As a result, CVaR is a sub-additive, convex functional of the financial portfolio.

1.3.2 Convex Risk Measures

To overcome the restrictive properties of positive homogeneity and subadditivity, Föllmer and Schied (2002) and Frittelli and Rosazza Gianin (2002) independently extended the axiomatic study of risk measures by relaxing the subadditivity and positive homogeneity conditions with the weaker convexity requirement. This requirement allows to control the risk of a convex combination by the combination of individual risks, and hence establishes a more general class of convex risk measures. A map $L^\infty \to \mathbb{R}$ is a convex risk measure if it satisfies the following axioms:

Translation invariance: For all real-valued random variables X and all real numbers (i.e., cashes) α, we have $\rho(X + \alpha \cdot r) = \rho(X) - \alpha$.

Convexity: For any two real-valued random variables X and Y, we have

$$\rho(\lambda X + (1-\lambda)Y) \leq \lambda\rho(X) + (1-\lambda)\rho(Y), \quad \forall\lambda \in [0,1].$$

Normalization: $\rho(0) = 0$.

Monotonicity: For any two real-valued random variables X and Y such that $X \leq Y$, we have $\rho(Y) \leq \rho(X)$.

The only law invariant relevant convex and time consistent dynamic risk measure is the entropic risk measure (Kupper and Schachermayer, 2009). This also suggests, especially in the dynamic case, the enlargement of the class of convex risk measures, provided we maintain the principle that diversification should not increase the risk.

For financial and economic purposes, sometimes it is convenient to reverse signs and emphasize on the utility of a financial position X rather than on its risk. For this purpose, if ρ is a convex risk measure, then $\phi(X) = -\rho(X)$ is called a concave monetary utility functional. If ρ is a coherent risk measure, then ϕ is called a coherent monetary utility functional.

1.3.3 Quasiconvex Risk Measures

Cerreia-Vioglio et al. (2011) introduced the idea that "diversification should not increase the risk" is exactly expressed by the quasi-convexity requirement:

Quasi-convexity: $\rho(\lambda X + (1-\lambda)Y) \leq \max\{\rho(X), \rho(Y)\}, \quad \forall\lambda \in [0,1],$

envelope or equivalently:

$$\text{the lower level sets } \{X \in L^\infty : \rho(X) \leq c\}, \quad \forall c \in \mathbb{R}, \text{ are convex.}$$

In fact, quasiconvexity and cash additivity imply convexity.

1.3.4 Dual Representation of Risk Measures

Some natural questions are as follows. Why are risk measures important and useful, and what are their appealing features in mathematical finance? There are two sources of risk measurement: model uncertainty and distribution risk.

One of the most appealing and useful properties of convex risk measures is their robustness against model uncertainty. This was first investigated in (Artzner et al., 1999), (Föllmer and Schied, 2002) and (Frittelli and Rosazza Gianin, 2002) for static risk measures. Under certain regularity conditions, a convex risk measure is represented as a suitably modified worst expected loss over a whole class of probabilistic models. Föllmer and Schied (2011) provided a comprehensive presentation of the theory of static coherent and convex risk measures.

Every convex risk measure $\rho : L^\infty \to \mathbb{R}$ on $\mathcal{G}$ admits a dual representation in the form of

$$\rho(X) = \sup_{Q \in \mathcal{P}} [\mathrm{E}_Q(-X) - \alpha(Q)],$$

where $\mathcal{P}$ is the set of all probability measures on $(\Omega, \mathcal{F})$ such that $\mathrm{E}_Q(X)$ is well defined for every $X \in \mathcal{G}$ and $Q \in \mathcal{P}$, and $\alpha : \mathcal{P} \to \mathbb{R} \cup \{+\infty\}$ is a penalty function. Moreover, for any probability measure $Q \in \mathcal{P}$, Q is absolutely continuous to P ($Q \ll P$). As discussed in (Carr et al., 2001), (Larsen et al., 2005) and (Föllmer and Schied, 2008, 2011), the elements of $\mathcal{P}$ can be interpreted as possible probabilistic models, taken more or less seriously according to the size of the penalty $\alpha(Q)$. Therefore the value $\rho(X)$ is computed as the worst-case expectation taken over all models $Q \in \mathcal{P}$ and penalized by $\alpha(Q)$.

The theory of dual representation was initiated in statistics in Huber (1981) and then extended to risk management in (Artzner et al., 1999), (Delbaen, 2002), (Föllmer and Schied, 2002), and (Frittelli and Rosazza Gianin, 2002). In the dual representation theory of convex risk measures, the goal is to derive a representation like $\rho(X) = \sup_{Q\in\mathcal{P}}[\mathrm{E}_Q(-X) - \alpha(Q)]$ in a systematic manner. The general idea is to apply the convex duality method. For each $Q \in \mathcal{P}$, define the minimal penalty function of ρ by $\alpha_\rho(Q) := \sup_{X\in\mathcal{G}}[\mathrm{E}_Q(-X) - \rho(X)] = \sup_{X\in\mathcal{A}_\rho} \mathrm{E}_Q(-X)$. With additional assumptions on the structure of $\mathcal{G}$ and on continuity properties of ρ, it is possible to derive the representation $\rho(X) = \sup_{Q\in\mathcal{P}}[\mathrm{E}_Q(-X) - \alpha_\rho(Q)]$ via the Fenchel-Legendre duality. In this case, the risk measure ρ is coherent if and only if α_ρ takes only the values 0 and $+\infty$. Therefore it arrives to the representation $\rho(X) = \sup_{Q\in\mathcal{Q}_\rho} \mathrm{E}(-X)$ where $\mathcal{Q}_\rho$ is the set of of all $Q \in \mathcal{P}$ such that $\alpha_\rho(Q) = 0$.

1.4 Deviation Measures

One of our goals in this thesis is to consider the concept of generalized deviation which extends the notion of standard deviation. Let $(\Omega, \mathcal{F}, P)$ be a probability space and L^2 be the space of square integrable functions on the probability space. To generalize the concept of standard deviation, we investigate certain functionals $\mathcal{D}$ on L^2 that satisfy the essential properties of standard deviation. The L^2-norm of the space L^2 is defined as $\|X\| = \sqrt{\mathrm{E}(X^2)}$.

According to (Rockafellar et al., 2006) and (Sarykalin et al., 2008), a functional $\mathcal{D} : L^2 \to [0, \infty]$ is called a deviation measure if it satisfies

(D1) $\mathcal{D}(C) = 0$ for any constant C, and $\mathcal{D}(X) > 0$ for any non-constant X,

(D2) $\mathcal{D}(\lambda X) = \lambda \mathcal{D}(X)$ for any random variable X and constant $\lambda > 0$ (positive homogeneity),

(D3) $\mathcal{D}(X + Y) \leq \mathcal{D}(X) + \mathcal{D}(Y)$ for any X and Y (subadditivity).

A deviation measure is called a lower range dominated deviation measure if it satisfies the axiom

(D4) $\mathcal{D}(X) \leq \mathrm{E}\, X + \sup(-X)$ for all $X \in L^2$.

The standard deviation $\sigma(X) = \|X - \mathrm{E}\, X\|$ clearly satisfies all these properties. This definition of a generalized deviation measure allows $\mathcal{D}(X) = \infty$ for some random variable X. When this is excluded, we have a finite deviation measure. In (D2) and (D3), infinite values are handled in the common ways: $\alpha + \infty = \infty$ for any $\alpha \in (-\infty, \infty]$, $\lambda\infty = \infty$ for any $\lambda > 0$, and $0\infty = 0$. The combination of properties (D2) and (D3) implies that $\mathcal{D}$ is a convex functional on L^2. Furthermore,

$\mathcal{D}(X)$ depends only on $X - \mathrm{E}\,X$ and vanishes only when $X - \mathrm{E}\,X = 0$ almost surely. This captures the idea that $\mathcal{D}$ measures the degree of uncertainty in X. Indeed, $\mathcal{D}$ acts as a sort of norm on the pure uncertainty subspace of L^2 consisting of all X with $\mathrm{E}\,X = 0$. Note that the symmetry property of standard deviation $\sigma(X)$ (that is, $\sigma(X) = \sigma(-X)$) may not be satisfied by a generalized deviation measure. However, if $\mathcal{D}$ is a deviation measure, then so too are its reflection $\tilde{\mathcal{D}}(X) = \mathcal{D}(-X)$ and its symmetrization $\hat{\mathcal{D}}(X) = \frac{1}{2}\left[\mathcal{D}(X) + \tilde{\mathcal{D}}(X)\right]$.

Deviation measures are extension of the concept of standard deviation and can be constructed based on other measures of variability of a random variable.

Example 1. (Semideviation measures) The first kind of variability based on which generalized deviation measures may be introduced is the deviation from the mean $X - \mathrm{E}\,X$. The standard deviation $\sigma(X) = \|X - \mathrm{E}\,X\|$ is one example. The standard upper and lower semideviation measures $\sigma_+(X) = \|(X - \mathrm{E}\,X)^+\|$ and $\sigma_-(X) = \|(X - \mathrm{E}\,X)^-\|$ satisfy all the properties (D1) - (D4) but are not symmetric. All these deviation and semideviation measures $\sigma(X)$, $\sigma_+(X)$ and $\sigma_-(X)$ are finite and continuous.

Example 2. (Range-based generalized deviation measures) More generalized deviation measures may be introduced based on the range. A deviation measure that is lower semicontinuous is given by $\mathcal{D}_l(X) = \mathrm{E}\,X - \inf X = \sup[\mathrm{E}\,X - X]$, which measures the size of the lower range of X. It reflection, $\mathcal{D}_u(X) = \tilde{\mathcal{D}}_l(X) = \sup X - \mathrm{E}\,X = \sup[X - \mathrm{E}\,X]$ is also lower semicontinuous, but measures the size of the upper range of X. Their sum, measuring the full range of X, $\mathcal{D}(X) = \mathcal{D}_l(X) + \mathcal{D}_u(X) = \sup X - \inf X$ is also semicontinuous. Unless the probability space is essentially finite, these deviation measures can take on ∞ and need not be continuous.

Deviation measures have several important properties. The following proposition talks about their continuity.

Proposition 1.4.1 (Continuity of deviation measures (Rockafellar et al., 2006)). A finite deviation measure $\mathcal{D}$ on L^2 that is lower semicontinuous must be continuous.

Another important property of the generalized deviation measure is given by the concept of dominance, which has a connection with duality. When a random variable X is used to measure the wealth and health of a financial institution, low values of X call for special attention. The lower range dominance property (D4) is clearly directed toward particular concerns about outcomes of X possibly falling short of $\mathrm{E}\,X$. The next result demonstrates that this concept supports a probabilistic interpretation of $\mathcal{D}$ in terms of downside risk. Among the generalized deviation measures we have just introduced, the lower range dominance holds for $R_l(X)$ and $\sigma_-(X)$, but not for $R_u(X)$, $\sigma(X)$ or $\sigma_+(X)$.

Theorem 1.4.1 (Dual characterization of deviations (Rockafellar et al., 2006)). A functional $\mathcal{D} : L^2 \to [0, \infty)$ is a lower semicontinuous deviation measure if and only if it has a representation of the form

$$\mathcal{D}(X) = \mathrm{E}\, X + \sup_{L^Q \in \mathcal{Q}} \mathrm{E}[-XL^Q], \tag{1.5}$$

where $\mathcal{Q}$ is a subset of L^2 having the following properties:

(Q1) $\mathcal{Q}$ is nonempty, closed and convex,

(Q2) for every nonconstant X there is some $L^Q \in \mathcal{Q}$ with $\mathrm{E}[XL^Q] < \mathrm{E}\, X$,

(Q3) $\mathrm{E}\left[L^Q\right] = 1$ for all $L^Q \in \mathcal{Q}$.

In this representation, $\mathcal{Q}$ is uniquely determined by $\mathcal{D}$ through

$$\mathcal{Q} = \{L^Q | \mathcal{D}(X) \geq \mathrm{E}\, X - \mathrm{E}[XL^Q] \text{ for all } X\}, \tag{1.6}$$

and the finiteness of $\mathcal{D}$ is equivalent to the boundedness of $\mathcal{Q}$.

Furthermore, $\mathcal{D}$ is lower range dominated if and only if $\mathcal{Q}$ has the additional property that

(Q4) $L^Q \geq 0$ for $L^Q \in \mathcal{Q}$.

Clearly for a lower semicontinuous deviation measure $\mathcal{D}$, the set $\mathcal{Q}$ plays a special role. From now on, we focus on L^Q which satisfies the conditions Q3 and Q4. The densityL^Q can be regarded as the density relative to P of some probability measure Q on Ω. Equivalently, such measures Q have a well defined density $L^Q = \dfrac{dQ}{dP} \in L^2(\Omega, \mathcal{F}, P)$. When L^Q is the density for Q, the expectation of a random variable X with respect to Q instead of P is

$$\mathrm{E}\left[\frac{dQ}{dP}X\right] = \int_{\Omega} X(\omega)\frac{dQ}{dP}(\omega)dP(\omega) = \int_{\Omega} X(\omega)dQ(\omega) = \mathrm{E}_Q[X]. \tag{1.7}$$

The set of all probability density functions on the probability space $(\Omega, \mathcal{F}, P)$ can be written as

$$\mathcal{M} = \{L^Q \in L^2 | L^Q \geq 0, \mathrm{E}\left[L^Q\right] = 1\}$$

In contemplating a subset $\mathcal{Q}$ of $\mathcal{M}$, one is essentially looking at some collection of alternatives Q to P. This could be motivated by a reluctance to accept P as furnishing a completely reliable model of the relative occurrences of the future states and the desire to test the dangers of two much trust in P. A subset $\mathcal{Q} \in \mathcal{M}$ is a risk envelope if it is nonempty, closed and convex. The uniquely determined set $\mathcal{Q}$ in Theorem 1.4.1 is called the risk envelope of the lower semicontinuous deviation measure $\mathcal{D}$.

Rockafellar (2007) also defined the risk envelope for the coherent risk measure and provided some examples.

Theorem 1.4.2 (Rockafellar, 2007)**.** There is a one-to-one correspondence $\rho \leftrightarrow \mathcal{Q}$ between the coherent risk measure ρ and the risk envelope $\mathcal{Q}$, given by the relations

$$\rho(X) = \sup_{Q \in \mathcal{Q}} \mathrm{E}_Q[-X], \text{ where } \mathcal{Q} = \{L^Q | \mathrm{E}_Q[-X] \leq \rho(X) \text{ for all } X\}.$$

Some examples are

$$\rho(X) = \mathrm{E}[-X] \leftrightarrow \mathcal{Q} = \left\{L^Q \in \mathcal{M} | L^Q = 1\right\},$$
$$\rho(X) = \sup[-X] \leftrightarrow \mathcal{Q} = \mathcal{M},$$
$$\rho(X) = \mathrm{CVaR}_\alpha(X) \leftrightarrow \mathcal{Q} = \left\{L^Q \in \mathcal{M} | L^Q \leq \alpha^{-1}\right\}.$$

We can think of these probability measures as designating alternatives to the underlying probability measure P. The difference $\mathrm{E}\,X - \mathrm{E}_Q[X]$ assesses how much worse the expectation of X might get under Q than under P. In this sense, $\mathcal{D}$ performs a worst-case analysis over the probability alternatives that have been selected. It can be derived from (1.5) that

$$\mathcal{D}(X) + \tilde{\mathcal{D}}(X) = \sup_{Q \in \mathcal{Q}} \mathrm{E}_Q[X] - \inf_{Q \in \mathcal{Q}} \mathrm{E}_Q[X],$$

which measures the difference between the best and worst possible expectations for X relative to the risk envelope $\mathcal{Q}$. This means that $\tilde{\mathcal{D}}$ has a similar form as well.

A wide range of examples will emerge in due course, but the most extreme case of a lower range dominated deviation measure can immediately be seen as associated with the largest possible set $\mathcal{Q}$ satisfying (Q1) - (Q4):

$$\mathcal{D}(X) = \mathrm{E}\,X - \inf X, \quad \text{corresponds to} \quad \mathcal{Q} = \mathcal{M}. \tag{1.8}$$

This deviation measure performs a worst-case analysis of $\mathrm{E}\,X - \mathrm{E}_Q[X]$ by looking at all the possible alternatives P' to P in this scheme. In contrast,

$$\mathcal{D}(X) = \sigma(X), \quad \text{corresponds to} \quad \mathcal{Q} = \left\{Q \left| \sigma\left(1 - L^Q\right) \leq 1, \mathrm{E}\left[L^Q\right] = 1\right.\right\}. \tag{1.9}$$

In this case the elements Q of $\mathcal{Q}$ fail to necessarily satisfy $L^Q \geq 0$, as is consonant with the lack of lower range dominance by $\mathcal{D} = \sigma$.

The probability measure P itself corresponds to the density function $\dfrac{dP}{dP} = 1$. From (1.6), every risk envelope $\mathcal{Q}$ has P as one of its elements. Thus, in the probabilistic interpretation when $\mathcal{Q}$, associated with a lower semicontinuous deviation measure $\mathcal{D}$, satisfies (Q4), the specified collection of probability measures $\mathcal{Q}$ must in particular include P. In fact, through (Q2) it must constitute a kind of neighbourhood of P in some sense.

It is worth noting that, although (Q1) is essential in establishing the one-to-one correspondence between the functional $\mathcal{D}$ and its envelope $\mathcal{Q}$ in Theorem 1.4.1, formula (1.5) would still define a lower semicontinuous deviation measure $\mathcal{D}$ if $\mathcal{Q}$ merely satisfies (Q2) and (Q3). The risk envelope associated with $\mathcal{D}$ is then the closure of the convex hull of $\mathcal{Q}$. Moreover, for every $B \in \mathcal{F}$,

$$Q(B) = \int_B \frac{dQ}{dP}, P(B) = \int_B 1dP.$$

Hence

$$\mathrm{E}\left[\frac{dQ}{dP}\right] = 1 \Rightarrow Q(B) = P(B), \quad \forall B \in \mathcal{F}.$$

This means property (Q3) implies that Q and P are equivalent.

1.4.1 Relation to Coherent Risk Measures

Although deviation measures are introduced for applications involving risk, they are not "risk measures" in the sense proposed in the landmark paper by Artzner et al. (1999). The connection between deviation measures and risk measures is strong, but there is a crucial distinction. Instead of measuring the uncertainty in X, a risk measure evaluates the "overall seriousness of possible losses" associated with X. A loss is an outcome below 0, in contrast to a gain which is an outcome above 0. When applying a risk measure, the orientation is crucial. If the concern is over the extent to which a given random variable X dropping below a threshold C, one needs to replace X by $X - C$. Moreover, a deviation measure is always positive but a risk measure can be either positive or negative.

The central notion for this section is the following concept of a "coherent" risk measure, adopted from (Artzner et al., 1999).

Definition 1.4.1 (Coherent risk measures)**.** By a coherent risk measure, we mean any functional $\rho : L^2 \to (-\infty, \infty]$ satisfying

(R1) $\rho(X + C) = \rho(X) - C$ for any X and any C,

(R2) $\rho(0) = 0$, and $\rho(\lambda X) = \lambda\rho(X)$ for any X and any $\lambda > 0$,

(R3) $\rho(X + Y) \leq \rho(X) + \rho(Y)$ for any X and Y,

(R4) $\rho(X) \leq \rho(Y)$ when $X \geq Y$,

where X and Y are random variables, and C and λ are constants.

The rationale behind coherent risk measures was well argued in (Artzner et al., 1999). However, how are risk measures related to the deviation measures we have introduced? Although (R2) and (R3) for a risk measure agree with (D2) and (D3) for a deviation measure, (R1) and (D1) are entirely different, in fact mutually incompatible - no functional on L^2 can satisfy both (R1) and (D1) simultaneously. Despite this, there is a simple relationship between the two kinds of measures. In order to explain this clearly, let us introduce a slightly different concept of risk measure.

Definition 1.4.2 (Strict expectation-boundedness)**.** A risk measure $\rho : L^2 \to (-\infty, \infty]$ is strictly expectation-bounded if it satisfies the properties (R1), (R2) and (R3) of Definition 1.4.1, and

(R5) $\rho(X) > \mathrm{E}[-X]$ for all nonconstant X .

When all properties (R1), (R2), (R3), (R4) and (R5) are satisfied, we speak of a coherent, strictly expectation-bounded risk measure.

The strict inequality in (R5) is key, since (R1) already guarantees that $\rho(X) = \mathrm{E}[-X]$ when X is a constant random variable. This is why we speak of "strict expectation-boundedness". The version with a weak inequality in place of the strict inequality would accordingly be "expectation-boundedness". Neither version was contemplated in (Artzner et al., 1999), where no reference probability distribution was assigned to the state space Ω.

Theorem 1.4.3 (Strictly expectation-bounded risk measures)**.** There is a one-to-one correspondence between the deviation measure $\mathcal{D}(X)$ and the strictly expectation-bounded risk measure $\rho(X)$, given by the relationships

(a) $\mathcal{D}(X) = \rho(X - \mathrm{E}\,X)$,

(b) $\rho(X) = \mathrm{E}[-X] + \mathcal{D}(X)$.

Specifically, if ρ is a strictly expectation-bounded risk measure and $\mathcal{D}$ is defined by (a), then $\mathcal{D}$ is a deviation measure that yields back ρ through (b). On the other hand, if $\mathcal{D}$ is a deviation measure and ρ is defined by (b), then ρ is a strictly expectation-bounded risk measure that yields back $\mathcal{D}$ through (a). In this correspondence, the risk envelope $\mathcal{Q}$ associated with $\mathcal{D}$ (in the presence of lower semicontinuity) furnishes for ρ the representation

$$\rho(X) = \sup_{Q \in \mathcal{Q}} \mathrm{E}_Q[-X]. \tag{1.10}$$

Furthermore, ρ is coherent if and only if $\mathcal{D}$ is lower range dominated. Thus, lower range bounded deviation measures (satisfying (D1) - (D5)) correspond one-to-one to coherent, strictly expectation-bounded risk measures (satisfying (R1) - (R5)) under the relationships (a) and (b).

Under the correspondences (a) and (b) in Theorem 1.4.3, $\mathcal{D}$ is called the deviation measure associated with ρ, whereas ρ is called the risk measure associated with $\mathcal{D}$. Note that ρ is finite if and only if $\mathcal{D}$ is finite. Likewise, ρ is lower semicontinuous (or continuous) if and only if $\mathcal{D}$ is lower semicontinuous (or continuous). Moreover, ρ is homogeneous if and only if $\mathcal{D}$ is homogeneous.

For coherent risk measures, the representation in (1.10) has an interpretation similar to that in (1.5) for deviation measures. The elements $Q \in \mathcal{Q}$ can be viewed as alternative probability measures of P. The quantity $\rho(X) = \sup_{Q\in\mathcal{Q}} \mathrm{E}_Q[-X]$ then designates the loss under that alternative, and ρ identifies the worst possible loss with respect to the specified class of alternatives.

Example 1. (CVaR-deviation) For any $\lambda \in (0,1)$, the functional

$$\mathrm{CVaRD}_\lambda(X) = \mathrm{CVaR}_\lambda(X - \mathrm{E}\, X)$$

is a continuous, lower range dominated deviation measure on L^2 for which the risk envelope is

$$\mathcal{Q} = \left\{ L^Q \in \mathcal{M} | L^Q \leq \frac{1}{\lambda} \right\}.$$

This corresponds to the coherent, strictly expectation-bounded risk measure $\mathrm{CVaR}_\lambda(X)$.

1.5 Summary and Structure of the Thesis

In this chapter, we have reviewed the history and some background materials in mathematical finance. The materials relevant to our thesis research include those on risk measures, deviation measures and their generalizations. The useful models of Markov decision processes and partially observable Markov decision processes are also reviewed.

The PhD thesis is structured as follows. In Chapter 2, we extend the classic model of Markov decision processes simultaneously in two directions: partially observable Markov decision processes and risk sensitive Markov decision processes.

That is, we introduce and solve the model of risk sensitive partially observable Markov decision processes.

In Chapter 3, we discuss conditional risk and deviation measures, and introduce a new type of conditional deviation measures and define the dynamic deviation measures.

In Chapter 4, using the model of Markov decision processes, we characterize and derive new classes of dynamic risk and deviation measures.

Chapter 5 concludes the thesis by summarizing achieved results and highlighting future research problems.

Chapter 2

Risk Sensitive Control of Partially Observable Markov Decision Processes

2.1 Introduction

Since 1960's, Markov decision processes have been a main stream research in operations research, with widespread and significant applications in many areas such as finance and business, industry and agriculture, medicine and clinical trials and many more. As the combination of the stochastic model of Markov processes and the technique of dynamic programming, Markov decision processes are essentially controlled Markov processes, in that a controller is able to influence the dynamics of the underlying Markov processes for the purpose of achieving a certain performance criterion.

Markov decision processes have been generalized in many directions. The most notable extensions are partially observable Markov decision processes and risk sensitive Markov decision processes.

In this chapter, we combine these two extensions and discuss risk sensitive control of partially observable Markov decision processes.

2.2 Partially Observable Markov Decision Processes

Partially observable Markov decision processes (POMDPs) may be introduced in different ways. One way is to assume that only partial observations of the states $s_n, n = 1, 2, \cdots$, are available in the form of a stochastic process $x_n, n = 1, 2, \cdots$, taking values in the observation space X. The unobservable states and observation processes are related by the probability distributions

$$P(x_{n+1}|x_1, \cdots, x_n; s_1, \cdots, s_{n+1}; a_1, \cdots, a_n) = P(x_{n+1}|s_{n+1}; a_n).$$

The controlled Markov process starts at initial state s_1 following a distribution p_1, then the first observation x_1 becomes available, and a decision a_1 is taken. The process makes a transition to state s_2 at the next stage and the second observation x_2 is available, and a decision a_2 is taken. At time n, the process is at state s_n and x_n is observed, a decision a_n is taken, and the process moves to state s_{n+1} and x_{n+1} is observed. The process repeats. Given a distribution p_1 of the initial state s_1 of the controlled Markov process, the state transition probabilities $P(s_{n+1}|s_n, a_n)$ and the observation distributions $P(x_{n+1}|s_{n+1}; a_n)$, the strategy π defines a probability measure $P_{p_1}^{\pi}$ on the random sample path

$(s_1, x_1, a_1, s_2, x_2, a_2, \cdots)$, which in turn defines an expectation $E^{\pi}_{p_1}$. With this expectation, the optimality criteria for regular Markov decision processes (MDP) are applicable to POMDPs. A standard approach to the POMDPs model is to formulate it as an equivalent completely observable MDP with the same action space but a new Borel state space consisting of probability distributions on the original state space.

Another way to introduce partially observable Markov decision processes is to assume that the transition probabilities $P(s_{n+1}|s_n, a_n)$ contain unknown parameters. There are certainly different statistical methods dealing with the unknown parameters. The Bayesian method is notably interesting and useful, because the process from the prior distribution to the posterior distribution is a natural process of information gathering. Information gathering is essential in many sequential decision problems, such as Markov decision processes.

In many cases, the advantage of partially observable Markov decision processes over the traditional Markov decision processes is that the partially observable models allow for even more forms of uncertainty in the underlying Markov processes. The disadvantage is that the computational complexity of algorithms for optimally solving the partially observable modes are extremely difficult.

The study of partially observable Markov decision processes has a long history. Åström (1965) appeared to be one of the earliest works assuming incomplete state information. This is followed by many important works, including (Smallwood and Sondik, 1973). White III (1991) provided a survey of solution techniques for the partially observable Markov decision processes. Fernández-Gauchercurd et al. (1991) investigated the average criteria and structured properties for partially observable Markov decision processes. Lovejoy (1991) provided a survey for the algorithm for solving partially observable Markov decision processes.

Suppose that at state s_t, action a_t is taken with a reward $r_t(s_t, a_t)$. The state s_t is unobservable and its information state is b_t, which is a posterior distribution of the underlying state s_t given the past history (up to time $t-1$) of the observations. Let S be the state space, then b_t is a distribution over S, and we denote $b_t(s)$ as the probability of $s \in S$. Denote the state-transition distribution by $P_{\text{tran}}(s'|s, a)$, which is the probability of making a transition to s' at the next stage, given the state s and action a at the current stage. Also denote the observation law $P_{\text{obs}}(x|s, a)$, which gives the probability of observing the signal x when the state is s and action is a at the current stage. Then the next information state given current state s_t and a_t is given by the following two-step procedure, commonly named a Kalman filter for a particle filter (Ristic et al., 2004).

Step1. Use Bayes' formula to update the belief state $\hat{b}_t$ using the signal x_t:

$$\hat{b}_t(s_t) = \frac{P_{obs}(x_t|s_t, a_t)b_t(s_t)}{\sum_{s_t \in S} P_{obs}(x_t|s_t, a_t)b_t(s_t)}, \qquad s_t \in S.$$

Step 2. Derive the information state at the next stage using the state-transition law

$$b_{t+1}(s_{t+1}) = \sum_{s_t \in S} \hat{b}_t(s_t) P_{\text{tran}}(s_{t+1}|s_t, a_t), \quad s_{t+1} \in S.$$

Under this formulation, the partially observable Markov decision process becomes a completely observable Markov decision process (COMDP) whose state is the information state taking values in the space of all distributions over the original state space S Chong et al. (2009). Note that this approach is easily extended to the case where the state space, signal space and action space are all Borel spaces, by replacing summation with integration.

In the process of reformulating a POMDP into a COMDP, the original unobservable state s_t is replaced by its distribution b_t. To investigate the COMDP, it is important to understand the structure of the set of all distributions.

Let S be the original state space and $\mathcal{B}(S)$ be the set of all distributions on S. By Parthasarathy (1967), it is well known that if S is a Borel space, i.e. a Borel subset of a Polish space (complete, separable metric space), then $\mathcal{B}(S)$ is also a Polish space with respect to the topology given by the weak convergence of probability measures.

Partially observable Markov decision processes have been applied to several areas. Tomberlin (2010) applied the model of partially observable Markov decision processes to optimally manage the potential marbled murrelet nesting sites. The marbled murreler (Brachyramphus marmoratus) is an endangered species of seabird that spends most of its life at sea and nests in costal forests from Alaska to California. Goulionis and Vozikis (2012) used the model of partially observable Markov decision processes to solve the problem of optimal management of patients, who are characterized by hidden disease states and medical procedures. Nakai (2010) applied the model of partially observable Markov decision processes to develop an optimal maintenance policy for some products, during whose life cycle, a condition changes and causes some trouble. Vozikis et al. (2012) applied a special form of partially observable Markov decision processes to determine an optimal strategy for the treatment of patients with ischemic heart disease. Brooks et al. (2006) applied a continuous state partially observable Markov decision process to investigate the optimization problem of decision-making under uncertainty for robot navigation. Chong et al. (2009) applied partially observable Markov decision process to adaptive sensing. Agussurja and Lau (2011) applied partially observable Markov decision processes to investigate the optimization problem of modelling and guiding taxi cruising in a congested city which consists of a large number of taxis and passengers. Arifoğlu and Özekici (2010) used the model of partially observable Markov decision processes to analyze a single-item periodically reviewed inventory system with random yield and finite capacity which is operated in a random environment.

2.3 Risk Sensitive Markov Decision Processes

The traditional study of Markov decision processes are risk neutral in that the rewards are additive. However we know that our attitude towards risk will affect our decisions. In this section, we look at one extension of Markov decision processes: risk sensitive Markov decision processes (RSMDPs).

Studies of RSMDPs started with the seminal paper by Howard and Mathesons (1972), who used the exponential utility function. Their idea was to adjust the transition probabilities which depends only on the current state, so that a kind of recursive equations might still hold. Kreps (1977a,b) extended the approach by allowing the utility function to depend on the complete history of the process. Chung and Sobel (1987) considered a more general optimization criteria and White (1988) provided a review on the use of other probabilistic criteria.

As another extension, risk measures such as value at risk may be incorporated into the utility function, see for example Ruszczyński (2010).

One approach to RSMDPs is to optimize the certainty equivalent of the stochastic optimization criterion. Bäuerle and Riedel (2014) continued the use of certainty equivalent and risk sensitive criteria. They assumed a general utility function and applied the ideas in Kreps (1977a,b) to expand the state space in order to use the technique of dynamic programming.

In this chapter, we extend the results in Bäuerle and Riedel (2014) to partially observable Markov decision process. That is, we work on the combination of RSMDPs and POMDPs to generate risk sensitive partially observable Markov decision processes (RSPOMDPs).

Let U be the utility function which is continuous and strictly increasing so that the inverse U^{-1} exists. Let X be a random variable. Then the certainty equivalent of X is defined to be a constant a such that $U(a) = E[U(X)]$. Therefore,

$a = U^{-1}(E[U(X)])$. Using Taylor expansion, it can be shown that $U^{-1}(\mathrm{E}[U(X)])$ is approximated by

$$U^{-1}(\mathrm{E}(U(X))) \approx \mathrm{E}[X] + \frac{1}{2}\frac{U''(\mathrm{E}X)}{U'(\mathrm{E}X)}\mathrm{Var}(X)$$

assuming that U is twice differentiable. This expansion makes use of both the mean and variance of the random variable X. Here, $-\dfrac{U''(x)}{U'(x)}$ is the Arrow-Pratt function of absolute risk aversion. Commonly used utility functions are the exponential function $U(x) = \frac{1}{\gamma}e^{\gamma x}, \gamma \neq 0$, and the power function $U(x) = \frac{1}{\gamma}x^{\gamma}, \gamma \neq 0$.

2.4 Risk Sensitive Control of Partially Observable Markov Decision Processes

Let b_0 be the initial information state which is a probability distribution of the unobservable initial state of the underlying controlled Markov process. Also let π be any strategy, our goal is to maximize

$$V_\pi(b_0) = \mathrm{E}_\pi[U(\sum_{t=0}^{T} r(s_t, a_t))|b_0] = \mathrm{E}_\pi[U(R^T(b_0)|b_0],$$

where $R^T(b_0) = \sum_{t=0}^{T} r(s_t, a_t)$. That is, we seek the optimal value $V(b_0) = \sup_\pi V_\pi(b_0)$ and an optimal strategy π^* such that $V(b_0) = V_{\pi^*}(b_0)$ for any $b_0 \in \mathcal{B}(S)$. Let r summarize the total award so far accumulated. Let $\mathbb{R}_+ = (0, \infty)$ be the space of all positive real numbers. Then we apply the technique of dynamic programming, and define, for each given distribution $b \in \mathcal{B}(S)$,

$$V_{t,\pi}(b, r) = \mathrm{E}_\pi[U(R^t(b) + r)], \quad b \in \mathcal{B}(S), r \in \mathbb{R}_+$$

and

$$V_t(b, r) = \sup_{\pi} V_{t,\pi}(b, r), \quad b \in \mathcal{B}(S), r \in \mathbb{R}_+.$$

This is the idea of formulating a bivariate Markov decision process.

Consider a partially observable Markov decision process with the following elements (S, X, A, P, Q, r):

- a Borel state space S, endowed with a Borel σ-algebra $\mathcal{S}$,
- a Borel signal space X, endowed with a Borel σ-algebra $\mathcal{X}$,
- a Borel action space A, endowed with a Borel σ-algebra $\mathcal{A}$,
- A Borel set $D \subset S \times A$ and subsets $D(s) = \{a \in A, (s, a) \in D\}$ of admissible actions in state s,
- the state transition law $P(ds'|s, a)$ which is a regular conditional distribution from D to S,
- the observation (or information) kernel $Q(dx|s, a)$, which is a regular conditional distribution from $S \times A$ to X,
- a measurable one-step reward function $r : D \to [r_1, r_2]$ with $0 < r_1 < r_2 < \infty$.

The evolution of a partially observable Markov decision process is as follows.

At the initial time $n = 0$, the initial unobservable state s_0 is given a prior distribution P_0. The initial information x_0 is observed according to the initial observation (or information) kernel $Q_0(dx|s_0)$. An initial action a_0 is taken from the admissible action set $D(s_0)$, with two consequences: first, the random initial reward $r(s_0, a_0)$ is collected; and second, the Markov process moves to a state s_1 at time $n = 1$ according to the state transition law $P(ds_1|s_0, a_0)$. Moreover, the information x_1 is generated at time $n = 1$ according to the observation (or information) kernel $Q(dx_1|a_0, s_1)$. This process then continues. The optimal criterion can be any of the optimality criteria listed for Markov decision process.

In order to ensure the existence of an optimal strategy, we follow Hernández-Lerma (1989) and assume that both the state transition law and the observation (or information) kernel are weakly continuous. We also assume a bounded and positive reward function, which is also continuous in its components. We assume that r_1 is positive in order to use the utility functions $U(x) = \ln x$ and $U(x) = \frac{1}{\gamma}x^\gamma$ with $\gamma < 0$. However for the exponential utility function $U(x) = \frac{1}{\gamma}e^{\gamma x}$, the lower bound is not necessary.

Although the computation of the optimal strategy for a POMDP is complex and much harder than that for an MDP, POMDPs can be solved in principle by reducing them to COMDPs, with the original state space being enlarged to the spaces of belief probabilities for the POMDPs (see (Sawaragi and Yoshikawa, 1970), (Rhenius, 1974), (Yushkevich, 1976), (Dynkin and Yushkevich, 1979), (Hernández-Lerma, 1989) and (Feinberg et al., 2012)).

For the risk sensitive control of the POMDPs, our goal is to maximize the certainty equivalents of the utilities of both finite horizon and infinite horizon criteria. Specifically, our objective is to maximize $U^{-1}(E_{b_0,\pi}[U(R^T)])$ or $U^{-1}(E_{b_0,\pi}[U(R_\alpha)])$ for any given initial distribution b_0 of the initial state, where the expectation E_π is taken with respect to the probability measure on the sample path generated by the strategy π, and $R^T = \sum_{t=0}^{T} r(s_t, a_t)$ is the finite horizon total reward and $R_\alpha = \sum_{t=0}^{\infty} \alpha^t r(s_t, a_t)$ is the infinite horizon total discounted reward.

To apply the technique of dynamic programming, we expand the state space as in (Bäuerle and Riedel, 2014). We formulate the risk sensitive POMDP as a bivariate MDP as follows.

- The new state space is $S^* = \mathcal{B}(S) \times \mathbb{R}_+$ given the product σ-algebra,
- The action space A is a Borel space with the Borel σ-algebra $\mathcal{A}$,
- The state transition law $P(ds'|s, a)$ is a regular conditional distribution from $S \times A$ to S,
- The reward function $r(s, a) : S \times A \to [r_1, r_2]$ is measurable.

Let Π be the space of all strategies. To apply the principle of dynamic programming, for each $t = 0, 1, \cdots, T$, define

$$V_{t,\pi}(b, r) = \mathrm{E}_{\pi,b}[U(R^t + r)], \quad b \in \mathcal{B}(S), r \in \mathbb{R}_+, \pi \in \Pi,$$

$$V_t(b, r) = \sup_{\pi} V_{t,\pi}(b, r), \quad b \in \mathcal{B}(S), r \in \mathbb{R}_+,$$

where r is the total reward so far accumulated by time t. The goal is to find the optimal value, for each given initial state $b_0 \in \mathcal{B}(S)$,

$$J_T(b_0) = \sup_{\pi \in \Pi} \mathrm{E}_{\pi,b_0}[U(R^T)] = \sup_{\pi \in \Pi} \mathrm{E}_{\pi,b_0}\left[U\left(\sum_{t=0}^{T} r(s_t, a_t)\right)\right], \quad b_0 \in \mathcal{B}(S)$$

and an optimal strategy π^* such that

$$J_T(b_0) = \mathrm{E}_{\pi^*,b_0}[U(R^T)], \quad b_0 \in \mathcal{B}(S).$$

Clearly, $V_T(b_0, 0) = J_T(b_0)$.

For the new MDP, over the state space $\mathcal{B}(S) \times \mathbb{R}_+$, the one stage reward is zero and the terminal reward is $V_0(b, r) = U(r)$. The state transition law is given by $\tilde{P}(db'|b, r, a)$ which is defined by

$$\int v(b', r')\tilde{P}(d(b', r')|b, r, a) = \int v(b', r(s, a) + r)\tilde{P}(db'|b, r, a)$$

where $v : \mathcal{B}(S) \times \mathbb{R}_+ \to \mathbb{R}_+$ is a bounded and continuous function. A deterministic strategy $\pi = (f_1, f_2, \cdots, f_T)$ is given by a sequence of measurable functions $f_i : \mathcal{B}(S) \times \mathbb{R}_+ \to A$. The strategy is Markov if each f_i depends only on the current state and not on the past history. The strategy is stationary if $f_i = f$ for all $i = 1, 2, \cdots, T$. Now, we define the space

$$\begin{aligned}
\mathcal{C} = \{v &: \mathcal{B}(S) \times \mathbb{R}_+ \to \mathbb{R}_+ \text{ such that } v \text{ is lower semicontinuous,} \\
&v(b, \cdot) \text{ is continuous and increasing for each given } b \in \mathcal{B}(S), \\
&\text{and } v(b, r) \geq U(r) \text{ for each } (b, r) \in \mathcal{B}(S) \times \mathbb{R}_+\}.
\end{aligned}$$

For each $v \in \mathcal{C}$ and each decision rule $f : \mathcal{B}(S) \times \mathbb{R}_+ \to A$, we define the operator

$$(T_f v)(b,r) = \int v(b', r(b, f(b,r)) + r) P(db'|b, f(b,r)), \quad (b,r) \in \mathcal{B}(S) \times \mathbb{R}_+.$$

The maximum reward operator is given by

$$(Tv)(b,r) = \sup_{a \in A} \int v(b', r(b,a) + r) P(db'|b,a), \quad (b,r) \in \mathcal{B}(S) \times \mathbb{R}_+.$$

If a decision function f achieves this maximum, that is, $T_f v = Tv$, then f is called a maximizer of v. Recall that $b \in \mathcal{B}(S)$ is the information state which gives a distribution over S. Therefore, we define

$$r(b,a) = \int_S r(s,a) db(s).$$

On the other hand, the transition $P(db'|b,a)$ is given by the Bayesian formula or the Kalman filter, as discussed under POMDPs.

Theorem 2.4.1. **(1)** For any strategy $\pi = (f_0, f_1, \cdots, f_T)$, we have the reward iteration $V_{t,\pi} = T_{f_0} \cdots T_{f_{t-1}} U$ for $t = 1, 2, \cdots, T$.

(2) The initial value is $V_0(b,r) = U(r)$ and the subsequent values are $V_t = TV_{t-1}$ for $t = 1, 2, \cdots, T$. That is,

$$V_t(b,r) = \sup_{a \in A} \int V_{t-1}(b', r(b,a) + r) P(db'|b,a).$$

Also, $V_t \in \mathcal{C}$ for each $t = 1, 2, \cdots, T$. Furthermore, for each $t = 1, 2, \cdots, T$, there exists a maximizer f_t^* of V_{t-1} and an optimal strategy $(g_0^*, \cdots, g_{T-1}^*)$ such that

$$g_t^*(h_n) = f_{T-t}^*(b_n, \sum_{k=1}^{t-1} r(b_k, a_k)), \quad t = 0, \cdots, T-1.$$

Proof. (1) We proceed by induction. The initial value is defined as $V_{0,\pi}(b,r) = U(r)$ for any strategy π. Now $V_{1,\pi}(b,r) = U(r(b, f_0(b,r)) + r) = (T_{f_0}U)(b,r)$. Suppose the statement is true for $V_{t-1,\pi}$, and let us consider $V_{t,\pi}$. We have

$$\begin{aligned}
&(T_{f_0} \cdots T_{f_{t-1}} U)(b,r) \\
&= \int V_{t-1,(f_1,f_2,\cdots)}(b', r(b, f_0(b,r)) + r) P(db'|b, f_0(b,r)) \\
&= \int E_{(f_1,f_2,\cdots),b'}[U(\sum_{k=0}^{t-2} r(b_k, a_k) + r(b, f_0(b,r)) + r)] P(db'|b, f_0(b,r)) \\
&= V_{t,\pi}(b,r).
\end{aligned}$$

The induction is completed.

(2) The results in part (1) show that the value functions in the reformulated model with bivariate states are the same value functions of the original model, therefore standard results on Markov decision processes may be applied. It is well known (Hinderer, 1970) that there is a Markov strategy that is optimal. On the other hand, Theorem 2.3.8 of Bäuerle and Riedel (2011) gives structure results that imply part (2), if we can show that $Tv \in \mathcal{C}$ and a maximizer for v exists, for each $v \in \mathcal{C}$.

Suppose that $v \in \mathcal{C}$. Then $(b, r, a, b') \to v(b', r(b,a) + r)$ is lower semicontinuous. Furthermore, $r \to v(b', r(b,a) + r)$ is both continuous and increasing. Then Theorem 17.11 of Hinderer (1970) implies that $(b, r, a) \to \int v(b, r, a, b') P(db'|b, a)$ is lower semicontinuous. Finally by proposition 2.4.3 in (Bäuerle and Riedel, 2011), the function $(b, r) \to \int (Tv)(b, r)$ is lower semicontinuous and there exists a maximizer of v.

By the Monotone Convergence Theorem, the function $r \to \int v(b', r(b,a) + r)P(db'|b,a)$ is both continuous and increasing, and hence lower semicontinuous. Because the supremum of any number of lower semicontinuous functions is again lower semicontinuous, the function $y \to (Tv)(b,r)$ is both continuous and increasing, and hence $(Tv)(b,r) \geq u(r)$. □

Applying the above theorem to the exponential utility function, we have the following results.

Corollary 2.4.1. For the exponential utility function $u(x) = \frac{1}{\gamma}e^{\gamma x}, \quad \gamma \neq 0$, we have

$$V_t(b,r) = e^{\gamma r}h_t(b), \quad t = 0, 1, \cdots, T \text{ and } J_T(b) = h_T(b),$$

where the function h_t are given by $h_0(b) = \frac{1}{\gamma}$ and

$$h_t(b) = \sup_{a \in A}\left\{e^{\gamma r(b,a)}\int h_{t-1}(b')P(db'|b,a)\right\}.$$

Proof. We demonstrate the results by induction. For $t = 0$, we have $V_0(b,r) = u(r) = \frac{1}{\gamma}e^{\gamma r}$ and so $h_0(b) = \frac{1}{\gamma}$.

Suppose the result is true for $t - 1$. Then by the principle of dynamic programming, we have

$$\begin{aligned}
V_t(b,r) &= \sup_a \int V_{t-1}(b', r(b,a)+r)P(db'|b,a) \\
&= \sup_a \int e^{\gamma(r(b,a)+r)} h_{t-1}(b')P(db'|b,a) \\
&= e^{\gamma r} \sup_a \{ e^{\gamma r(b,a)} \int h_{t-1}(b')P(db'|b,a)\} \\
&= e^{\gamma r} h_t(b).
\end{aligned}$$

The results follow. □

We now turn to the case of infinite horizon discounted criterion. The goal is to achieve

$$J_\infty(b_0) = \sup_{\pi\in\Pi} \mathrm{E}_{\pi,b_0}[U(R_\beta^\infty)] = \sup_{\pi\in\Pi} \mathrm{E}_{\pi,b_0}\left[U\left(\sum_{t=0}^{\infty} \beta^t r(s_t, a_t) \right) \right], \quad b_0 \in \mathcal{B}(S),$$

where $\beta \in (0,1)$ is the discount factor. The discount sequence is geometric.

For this problem, we need to expand the state space in order to keep track of the discounting sequence. The new state space is $\mathcal{B}(S) \times \mathbb{R}_+ \times (0,1)$, over which we define

$$\begin{aligned}
\mathcal{C} = \{ v : \mathcal{B}(S) \times \mathbb{R}_+ \times (0,1) \to \mathbb{R}_+ &\text{ such that } v \text{ is lower semicontinuous,} \\
v(b,\cdot,\cdot) &\text{ is continuous and increasing for each given } b \in \mathcal{B}(S), \\
\text{and } v(b,r,z) \geq u(r) &\text{ for each } (b,r,z) \in \mathcal{B}(S) \times \mathbb{R}_+ \times (0,1) \}.
\end{aligned}$$

Furthermore, a decision rule f is a measurable function from $\mathcal{B}(S) \times \mathbb{R}_+ \times (0,1)$ to A. For each $v \in \mathcal{C}$ and decision rule f, we define two operators

$$(T_f v)(b,r,z) = \int v(b', zr(b, f(b,r,z)) + r, z\beta) P(db'|b, f(b,r,z)),$$

where $(b, r, z) \in \mathcal{B}(S) \times \mathbb{R}_+ \times (0, 1)$, and

$$(Tv)(b, r, z) = \sup_a \int v(b', zr(b, a) + r, z\beta) P(db'|b, a), \quad (b, r, z) \in \mathcal{B}(S) \times \mathbb{R}_+ \times (0, 1).$$

To consider a risk-seeking decision maker, we assume that the utility function U is concave, increasing and differentiable. Let's define

$$V_{\infty,\pi}(b, r, z) = E_{b,\pi}[U(zR_\alpha^\infty + r)]$$

and

$$V_\infty(b, r, z) = \sup_{\pi \in \Pi} V_{\infty,\pi}(b, r, z), \quad (b, r, z) \in \mathcal{B}(S) \times \mathbb{R}_+ \times (0, 1)$$

Clearly, the optimal value is $J_\infty(b_0) = V_\infty(b_0, 0, 1)$ and can be derived recursively. A standard result for an infinite horizon Markov decision process is that there is a stationary strategy that is optimal. A stationary strategy can be written as $\pi = (f, f, \cdots) = f^\infty$. Recall that the reward r is assumed to be bounded between r_1 and r_2. Let's denote

$$\underline{b}(r, z) = U\left(z \frac{r_1}{1-\beta} + r\right)$$

and

$$\bar{b}(r, z) = U\left(z \frac{r_2}{1-\beta} + r\right).$$

Theorem 2.4.2. For the infinite horizon model, we have that V_∞ is the unique fixed point solution of $v = Tv$ in $\mathcal{C}$, and $\underline{b}(r, z) \leq v(b, r, z) \leq \bar{b}(r, z)$. Moreover, $T^n U$ increases to V_∞, $T^n \underline{b}$ increases to V_∞, and $T^n \bar{b}$ decreases to V_∞ as $n \to \infty$. Furthermore there is a maximizer f^* of V_∞ and $(g_0^*, g_1^*, \cdots)$ with $g_n^*(h_n) = f^*(b_n, \sum_{k=1}^{t-1} \beta^k r(b_k, a_k), \beta_n)$ being an optimal strategy.

Proof. First, let's introduce the finite horizon discounted criterion

$$J_{T,\beta}(b_0) = \sup_{\pi\in\Pi} \mathrm{E}_{b_0,\pi}[U(R^T_\beta)] = \sup_{\pi\in\Pi} \mathrm{E}_{b_0,\pi}[U(\sum_{k=0}^{T} \beta^k r(b_k, a_k))], \quad b_0 \in \mathcal{B}(S).$$

Furthermore, define

$$V_{t,\pi}(b, r, z) = \mathrm{E}_{b_0,\pi}[U(zR^t_\beta + r)], \quad (b, r, z) \in \mathcal{B}(S) \times \mathbb{R}_+ \times (0, 1)$$

and

$$V_t(b, r, z) = \sup_{\pi\in\Pi} V_{t,\pi}(b, r, z), \quad (b, r, z) \in \mathcal{B}(S) \times \mathbb{R}_+ \times (0, 1).$$

Then using the same idea as in the previous theorem, $V_t = T^t U$. Because of nonnegative rewards, the sequence V_t is increasing and $V_t \le V_\infty$ for all t.

We now show that $\lim_{t\to\infty} V_t = V_\infty$. Because $U : \mathbb{R}_+ \to \mathbb{R}_+$ is increasing, differentiable and concave, we have the inequality

$$U(v_1 + v_2) \le U(v_1) + U'(v_1)v_2, \quad v_1, v_2 \ge 0,$$

which can be verified by the Taylor expansion. Moreover, $U'(x) \ge 0$ and $U'(x)$ is decreasing. For any given state $(b, r, z) \in \mathcal{B}(S) \times \mathbb{R}_+ \times (0, 1)$ and any strategy π, we have

$$\begin{aligned}
V_{\infty,\pi}(b, r, z) &= \mathrm{E}_{b,\pi}[U(zR^\infty_\beta + r)] \\
&= \mathrm{E}_{b,\pi}[U(zR^\infty_\beta + r + \beta^t z \sum_{k=t}^{\infty} \beta^{k-t} r(s_k, a_k))] \\
&\le \mathrm{E}_{b,\pi}[U(zR^\infty_\beta + r)] + \beta^t z \frac{r_2}{1-\beta} \mathrm{E}_{b,\pi}[U'(zR^t_\beta + r)] \\
&\le V_{t,\pi}(b, r, z) + \beta^t z \frac{r_2}{1-\beta} U'\left(z(\frac{1-\beta^{t-1}}{1-\beta}) + r \right).
\end{aligned}$$

Write

$$\varepsilon_t = \beta^t z \frac{r_2}{1-\beta} U'\left(z(\frac{1-\beta^{t-1}}{1-\beta}) + r \right).$$

Then $\lim_{t\to\infty} \varepsilon_t = 0$, and we have $V_\infty \le \lim_{t\to\infty} V_t$. Combining with $V_t \le V_\infty$ for all t, we have that V_t increases to V_∞.

Because the operator T is increasing, we also have $TV_\infty \le T(V_t + \varepsilon_t) = V_{t+1} + \varepsilon_t$. Taking the limit when $t \to \infty$, we have $TV_\infty \le V_\infty$. On the other hand, $V_t \le V_\infty$ for all t implies $V_{t+1} = TV_t \le TV_\infty$ and hence $V_\infty \le TV_\infty$. Therefore, $V_\infty = TV_\infty$ and V_∞ is a fixed point of $v = Tv$. Now,

$$\begin{aligned} T\bar{b}(r,z) &= \sup_a U\left(\frac{z\beta r_2}{1-\beta} + zr(b,a) + r\right) \\ &\le U\left(\frac{z\beta r_2}{1-\beta} + zr_2 + r\right) \\ &= U\left(\frac{zr_2}{1-\beta}\right) = \bar{b}(r,z), \end{aligned}$$

and similarly $T\bar{b}(r,z) \ge \underline{b}(r,z)$. The operator T is increasing, so the sequence $T^n\bar{b}$ is decreasing, and the sequence $T^n\underline{b}$ is increasing and both sequences have limits.

Furthermore,

$$(T^t\underline{b})(b,r,z) = \sup_\pi \mathrm{E}_{b,\pi}\left[U\left(\frac{zr_1\beta^t}{1-\beta} + z\sum_{k=0}^{t-1} \beta^k r(s_k,a_k) + r \right)\right] \ge (T^tU)(b,r,z)$$

and $U(v_1+v_2) - U(v_1) \le U'(v_1)v_2$, we have

$$\begin{aligned} &0 \le (T^t\bar{b})(b,r,z) - (T^t\underline{b})(b,r,z) \\ &\le \sup_\pi \left\{ U\left(\frac{zr_2\beta^t}{1-\beta} + z\sum_{k=0}^{t-1}\beta^k r(s_k,a_k) + r\right) - U\left(z\sum_{k=0}^{t-1}\beta^k r(s_k,a_k) + r \right)\right\} \\ &\to 0 \text{ as } t \to \infty. \end{aligned}$$

Hence $\lim_{t\to\infty} T^t \underline{b} = \lim_{t\to\infty} T^t \bar{b}$. Together with $\underline{b} \le V_\infty \le \bar{b}$, we have $\lim_{t\to\infty} T^t \underline{b} = \lim_{t\to\infty} T^t \bar{b} = V_\infty$.

Finally, for the uniqueness of the fixed point, assume that $v = Tv$, with $\underline{b} \le v \le \bar{b}$. Then $T^t \underline{b} \le v \le T^t \bar{b}$. Taking the limit when $t \to \infty$ we have $v = V_\infty$. For the continuity of V_∞, note that it is lower semicontinuous because V_n is, and is upper semicontinuous because $T^t \bar{b}$ is. The existence of the maximizers follows from the continuity and compactness assumptions and standard results from the theory of Markov decision process. $\square$

2.5 Summary

The classical completely observed Markov decision processes (MDPs) have been generalized in many directions, including partially observable Markov decision processes (POMDPs), risk sensitive Markov decision processes (RSMDPs), and MDPs with risk measures.

In this chapter, we have combined the consideration of POMDPs and RSMDPs and extended MDPs further to risk sensitive partially observable Markov decision processes (RSPOMDPs). The main techniques used are to extend the RSPOMDPs model as a regular MDP by enlarging the state space and to apply dynamic programming.

Chapter 3

Extension of Static Risk and Deviation Measures

3.1 Introduction

Risk and deviation measures have been a very important area of research in mathematical finance. In fact they play fundamental roles in risk management because risk is an indispensable part of financial management.

Although very important, risk and deviation measures have been mostly studied in the static case. Especially for the deviation measure, there is no reference introducing the multi-period case formally. In this and the next chapters, we investigate two significant extensions of one period, static risk and deviation measures: conditional and dynamic risk measures, and conditional and dynamic deviation measures.

3.2 Extension of Static Risk Measures

Static risk measures can be extended in a number of different ways, most notably by assessing the risk situation in light of newly available information. One natural extension of static risk measures is to define conditional risk measures, which incorporate the information available at the time of risk assessment. Like the representation for a static risk measure as the worst expected loss over a whole class of probabilistic models, a conditional convex risk measure is represented as the worst conditional expected loss over a class of suitably penalized probability measures.

Along another direction, static risk measures are naturally extended to the dynamic setting. In this case, the sequence of information available is represented by a filtration and the sequence of dynamic risk measures is defined as a sequence of conditional risk measures adapted to the underlying filtration. Using stochastic models in the dynamical setting, dynamic risk measures are developed and updated over time in accordance with the new information (see (Riedel, 2004), (Frittelli and Rosazza Gianin, 2004), (Detlefsen and Scandolo, 2005), (Roorda et al., 2005) and (Cheridito et al., 2006)). One key tool used is the model of stochastic dynamic programming or equivalently, Markov decision processes.

3.2.1 Conditional Risk Measures

When additional information is provided, several authors, including Detlefsen and Scandolo (2005) and Föllmer and Penner (2006) defined the so-called conditional risk measures, because static risk measures do not incorporate information. For the given probability space $(\Omega, \mathcal{F}, P)$, consider a discrete-time multiperiod information structure $(\mathcal{F}_t)_{t=0,\cdots,T}$, which forms a filtration, with $\mathcal{F}_0 = \{\emptyset, \Omega\}$ and $\mathcal{F}_T = \mathcal{F}$. The set of all financial positions will be denoted as L^∞. Let L_t^∞ be the set of all $\mathcal{F}_t$-measurable P-a.s. bounded random variables. All inequalities and equalities applied to random variables are in the sense of almost surely.

A conditional convex risk measure is defined as follows.

Definition 3.2.1 (Föllmer and Penner, 2006)**.** A mapping $\rho_t : L^\infty \to L_t^\infty$ is called a conditional convex risk measure if it satisfies the following properties for all $X, Y \in L^\infty$:

Conditional translation invariance: For any $X_t \in L_t^\infty$,

$$\rho_t(X + X_t) = \rho_t(X) - X_t,$$

Monotonicity: $X \leq Y \Rightarrow \rho_t(X) \geq \rho_t(Y)$,

Conditional convexity: For all $\Lambda \in L_t^\infty$ with $0 \leq \Lambda \leq 1$,

$$\rho_t(\Lambda X + (1 - \Lambda)Y) \leq \Lambda\rho_t(X) + (1 - \Lambda)\rho_t(Y),$$

Normalization: $\rho_t(0) = 0$.

A conditional convex risk measure is called a conditional coherent risk measure if it has in addition the following property:

Conditional positive homogeneity: For all $\Lambda \in L_t^\infty$ with $\Lambda \geq 0$,

$$\rho_t(\Lambda X) = \Lambda \rho_t(X).$$

When $t = 0$, we obtain the usual definition of a convex risk measure. A weaker definition of a conditional convex is given in (Klöppel and Schweizer, 2007) where the normalization property is not required. In the dynamic setting it is possible to define risk measures for an income (or a payoff) stream, i.e. for stochastic processes instead of random variables, which we will discuss in the next chapter.

We denote $\mathcal{P}$ as the set of all probability measures on $(\Omega, \mathcal{F})$ which are absolutely continuous with respect to P, and $\mathcal{P}^e$ as the set of all probability measures on (Ω, F), which are equivalent to P on $\mathcal{F}$. Moreover, define the following two sets of probability measures as:

$$\mathcal{P}_t = \{Q \in \mathcal{P} | Q \sim P \text{ on } \mathcal{F}_t\} \tag{3.1}$$

$$\mathcal{Q}_t = \{Q \in \mathcal{P} | Q = P \text{ on } \mathcal{F}_t\}. \tag{3.2}$$

It is well known that an unconditional convex risk measure which is continuous from above has the representation

$$\rho(X) = \sup_{Q \in \mathcal{P}} \{\mathrm{E}_Q[-X] - \alpha(Q)\}, \quad X \in L^\infty,$$

in terms of a penalty function $\alpha : \mathcal{P} \to (-\infty, \infty]$. A similar characterization also holds for conditional convex risk measures which are continuous. In this more general representation, the expectations are conditional on the available information $\mathcal{F}_t$, the penalty function is random-valued and the supremum is in the essential sense. The penalty function will be given by a mapping from some set $\mathcal{Q} \subseteq \mathcal{P}_t$ to the set of $\mathcal{F}_t$-measurable random variable with values in $(-\infty, \infty]$ such that

$$\operatorname*{ess\,sup}_{Q \in \mathcal{Q}} \{-\alpha_t(Q)\} = 0.$$

We say that ρ_t has a robust representation if

$$\rho_t(X) = \operatorname*{ess\,sup}_{Q \in \mathcal{Q}} \{\mathrm{E}_Q[-X|\mathcal{F}_t] - \alpha_t(Q)\}, \quad \forall X \in L^\infty$$

with some set $\mathcal{Q} \subseteq \mathcal{P}_t$ and some penalty function α_t on $\mathcal{Q}$. In (Föllmer and Penner, 2006), there are two corollaries talking about the representation of conditional convex and coherent risk measures.

Corollary 3.2.1 (Föllmer and Penner, 2006)**.** We have the following representation results:

(a) A conditional convex risk measure ρ_t is continuous from above if and only if for any $P^* \in \mathcal{P}^e$ it is representable in the form

$$\rho_t(X) = \operatorname*{ess\,sup}_{Q \in \mathcal{Q}_t(P^*)} \{\mathrm{E}_Q[-X|\mathcal{F}_t] - \alpha_t(Q)\}, \quad X \in L^\infty,$$

where

$$\mathcal{Q}_t(P^*) := \{Q \in \mathcal{P} | Q = P^* \text{ on } \mathcal{F}_t, \mathrm{E}_{P^*}[\alpha_t(Q)] < \infty\}.$$

(b) A conditional coherent risk measure ρ_t is continuous from above if and only if for any $P^* \in \mathcal{P}^e$ it is representable in the form

$$\rho_t(X) = \operatorname*{ess\,sup}_{Q \in \mathcal{Q}_t^0(P^*)} \{\mathrm{E}_Q[-X|\mathcal{F}_t]\}, \quad X \in L^\infty,$$

where

$$\mathcal{Q}_t^0(P^*) := \{Q \in \mathcal{P} | Q = P^* \text{ on } \mathcal{F}_t, \alpha_t(Q) = 0 \ Q - a.s.\}.$$

The above corollary is quite general. In the rest of the thesis, we only consider the reference probability measure P instead of all probability measures $P^* \in \mathcal{P}^e$. Moreover, only the probability measures in $\mathcal{Q}_t$ can be chosen for the representation. This is due to several reasons (see (Detlefsen and Scandolo, 2005) and (Föllmer and Penner, 2006)). One important reason is that we need to use the natural equality $\mathrm{E}_P[\mathrm{E}_Q[X|\mathcal{F}_t]] = \mathrm{E}_P[X]$, which holds if and only if $Q \in \mathcal{Q}_t$.

One example of a conditional convex risk measure (but not coherent) with the explicit representation is the conditional entropic risk measure with the risk averse $\gamma > 0$, defined as

$$\rho_t^\gamma(X) = \frac{1}{\gamma} \log \mathrm{E}_P \left[e^{-\gamma X}|\mathcal{F}_t\right].$$

Conditional on the information $\mathcal{F}_t$, $\rho_t^\gamma(X)$ is the conditional capital requirement when the agent's attitude is described by the exponential utility $u_\gamma(x) = 1 - \exp(-\gamma x)$ for a risk aversion $\gamma > 0$.

For a parameter $\Lambda \in L_t^\infty$ with $0 < \Lambda < 1$ we consider the set of probabilistic models

$$\mathcal{Q}_\Lambda = \left\{ Q \in \mathcal{Q}_t \,\middle|\, \frac{dQ}{dP} \leq \Lambda^{-1} \right\}.$$

The corresponding conditional coherent risk measure

$$\mathrm{CVaR}_\Lambda(X|\mathcal{F}_t) = \operatorname*{ess\,sup}_{Q \in \mathcal{Q}_\Lambda} \{- \mathrm{E}_Q[X|\mathcal{F}_t]\}, \quad X \in L^\infty$$

is called the conditional CVaR at level Λ, which generalizes the unconditional risk measure $\mathrm{CVaR}_\lambda(X)$.

3.2.2 Dynamic Risk Measures

In this section we investigate conditional risk measures in a dynamic framework where successive measurements are performed.

Definition 3.2.2 (Detlefsen and Scandolo, 2005)**.** A dynamic risk measure is a family $(\rho_t)_{t=0,\cdots,T}$ such that $\rho_t : L^\infty \to L_t^\infty$ is a conditional risk measure. It is a dynamic convex (coherent) risk measure if all components ρ_t are conditional convex (coherent) risk measures.

A dynamic convex risk measure maps a random variable $X \in L^\infty$ into an adapted process $(\rho_t(X))_{t=0,\cdots,T}$, and can be seen as the result of the risk assessment of a final payoff through time. Definition (3.2.2) is quite general and it is interesting to look into further properties linking different components of a dynamic risk measure.

Time consistency: For any $X, Y \in L^\infty$ and $0 \le t \le T-1$,

$$\rho_{t+1}(X) = \rho_{t+1}(Y) \Rightarrow \rho_t(X) = \rho_t(Y).$$

Recursiveness: For any $X \in L^\infty$ and $0 \le t \le T-1$,

$$\rho_t(X) = \rho_t(-\rho_{t+1}(X)).$$

Supermartingale: For any $X \in L^\infty$ and $0 \le t \le T-1$,

$$\rho_t(X) \ge \mathrm{E}[\rho_{t+1}(X)|\mathcal{F}_t].$$

The financial meaning of time consistency is based on a general intuition: if two financial positions have the same risk tomorrow in every state of nature, then the same conclusion should be drawn today. The case for recursiveness, on the contrary, strongly relies on the validity of conditional translation invariance for the components ρ_t and hence on their interpretation as conditional capital requirement. In fact, if $\rho_{t+1}(X)$ is the conditional capital requirement that has to be set aside at time $t+1$ in view of the financial position X, then the risky position is equivalently described at time t by the amount $-\rho_{t+1}(X)$ occurring at t. Finally, the supermartingale property can be interpreted as follows: as time evolves, the information about the financial position X increases. This would lower the perceived riskiness - not almost surely, but in (conditional) mean. For a dynamic convex risk measure, time-consistency and recursiveness are equivalent. An example of a dynamic convex risk measure possessing all the above three properties is the dynamic entropic risk measure given by

$$\rho_t(X) = \frac{1}{\gamma} \log E\left[e^{-\gamma X} | \mathcal{F}_t\right], \quad X \in L^\infty$$

with a fixed risk aversion $\gamma > 0$.

Dynamic risk measures are natural extension of single period, static risk measures. However they must satisfy some important properties. One important property concerns about the interrelation of dynamic risk measures at different time points, referred to as time consistency, which was just discussed. Several notions of time consistency have been introduced in the literature. Strong time consistency, which corresponds to the principle of dynamic programming, is characterized by the additivity of acceptance sets and penalty functions, and also by a supermartingale property of the risk process and the penalty function process ((Frittelli and Rosazza Gianin, 2004), (Detlefsen and Scandolo, 2005), (Cheridito et al., 2006) and (Föllmer and Penner, 2006)).

Rejection consistency, also called prudence in (Penner, 2007), is a weaker version of time consistency. It ensures that one always stays on the safe side when updating risk assessment and so seems to be particularly suitable and reasonable from the point of view of a regulator. Weak acceptance and weak rejection consistency are the weakest notions of time consistency ((Weber, 2006), (Roorda et al., 2005), (Acciaio and Penner, 2011)). They require that if a certain financial position is accepted (or rejected) for any scenario tomorrow, it should be already accepted (or rejected) today.

As another important property, dynamic risk measures should be able to account for uncertainty about the time value of money (EI Karoui and Ravaelli, 2009). This leads to another extension of static risk measures by considering the entire cash flow process rather than the single total amount at the terminal date as the risky object. Dynamic risk measures for stochastic processes of financial states have been investigated ((Riedel, 2004) and (Frittelli and Scandolo, 2006)). For example, dynamic risk measures with a stochastic process of cash flows are considered in (Cheridito et al., 2006). However as pointed out in (Acciaio et al., 2010), dynamic risk measures for stochastic processes can be identified with risk measures for random variables on an appropriate product space. This allows a natural translation of results obtained for static risk measures for random variables to the process of dynamic measures.

3.3 Extension of Static Deviation Measures

3.3.1 Conditional Deviation Measures

In order to extend the one-to-one correspondence between the deviation measure $\mathcal{D}(X)$ and the strictly expectation-bounded risk measure $\rho(X)$ to the situation when additional information is provided, we need to extend the deviation measure to the conditional case. Similar to the techniques in (Föllmer and Penner, 2006), the additional information is described by a sub-σ-algebra $\mathcal{F}_t$ of the total information $\mathcal{F}$, and a conditional deviation measure is a mapping assigning to every square integrable $\mathcal{F}$-measurable random variable X, representing a financial position, a positive $\mathcal{F}_t$-measurable random variable $\mathcal{D}_t(X)$, representing the conditional variability of X.

A good feature of this space is their invariance with respect to the probability measure which is chosen in the equivalence class of P. Moreover, denote $(L_t^\infty)^+$ as the set of all positive $\mathcal{F}_t$-measurable P-a.s. bounded random variables on which the essential supremum can be taken. Moreover, the two sets of probability measures $\mathcal{P}_t$ and $\mathcal{Q}_t$ has been defined in Definition (3.1) and (3.2).

An element $X \in L^2$ describes a random financial position to be delivered to an agent at a fixed future time or the market value of a portfolio at that time. The σ-algebra $\mathcal{F}_t$ collects the information available to the agent who is assessing the riskiness of X. As a consequence the variability measurement of X leads to a positive random variable $\mathcal{D}_t(X)$ which is measurable with respect to $\mathcal{F}_t$, i.e., an element of the space $(L_t^\infty)^+$. Technically, a conditional deviation measure is a mapping $\mathcal{D}_t : L^2 \to (L_t^\infty)^+$. Plainly, we interpret $\mathcal{D}_t(X)(\omega)$ as the degree of variability of X when the state ω prevails. The σ-algebra $\mathcal{F}_t$ can be interpreted in different ways. It can model additional information available at time $t = 0$ to the agent. Alternatively, it can be interpreted as the information available at a future time $t > 0$, resulting from the observation of some variables related to the financial position X in the time interval $[0, t]$. In both cases, the sources of information can be public and shared by all agents, or private. Hence, conditional deviation measures open a way to the analysis of the consequences of asymmetric information for the measurement of variability.

Definition 3.3.1. A mapping $\mathcal{D}_t : L^2 \to (L_t^\infty)^+$ is called a conditional deviation measure if it satisfies the following properties for all $X, Y \in L^2$:

(D1') $\mathcal{D}_t(X) > 0$ and $\mathcal{D}_t(X_t) = 0$ for any $X_t \in (L_t^\infty)^+$.

(D2') Conditional positive homogeneity: For any $\Lambda \in (L_t^\infty)^+$, we have $\mathcal{D}_t(\Lambda X) = \Lambda \mathcal{D}_t(X)$.

(D3') Conditional subadditivity: $\mathcal{D}_t(X + Y) \leq \mathcal{D}_t(X) + \mathcal{D}_t(Y)$.

A conditional deviation measure is said to be a conditional lower range dominated deviation measure if it satisfies the first three properties above and

(D4') $\mathcal{D}_t(X) \leq \mathrm{E}[X|\mathcal{F}_t] + \sup(-X|\mathcal{F}_t)$.

We have $(L_0^\infty)^+ = [0, \infty]$ and obtain the usual definition of a deviation measure in (Rockafellar et al., 2006) when there is no initial information. The economic rationale behind the properties characterizing conditional deviation measures is the same as in the static case (see (Rockafellar et al., 2006), for instance).

Examples of conditional deviation measures can be naturally extended from the (static) deviation measures.

Example 1. (Conditional standard deviation)

$$\mathcal{D}_t(X) = (\mathrm{E}[(X - E[X|\mathcal{F}_t])^2|\mathcal{F}_t])^{1/2}.$$

Example 2. (Conditional range-based generalized deviation measures)

$$\mathcal{D}_t(X) = \mathrm{E}[X|\mathcal{F}_t] - \inf(X|\mathcal{F}_t) = \sup\{\mathrm{E}[X|\mathcal{F}_t] - X|\mathcal{F}_t\}.$$

Example 3. (Conditional CVaRD)

$$\mathrm{CVaRD}_\Lambda(X|\mathcal{F}_t) = \mathrm{CVaR}_\Lambda(X - \mathrm{E}[X|\mathcal{F}_t]|\mathcal{F}_t),$$

where $\mathrm{CVaR}_\Lambda(X|\mathcal{F}_t) = \inf_{Z\in L^2} \mathrm{E}[Z + \lambda^{-1}(X + Z)^-|\mathcal{F}_t]$ is the conditional CVaR.

Now we introduce the dual representation of the conditional CVaRD in the following theorem.

Theorem 3.3.1. On the set of probabilities

$$\mathcal{Q}_\Lambda = \left\{ Q \in \mathcal{Q}_t \,\middle|\, \frac{dQ}{dP} \le \Lambda^{-1} \right\}.$$

the conditional CVaRD has the representation

$$\mathrm{CVaRD}_\Lambda(X|\mathcal{F}_t) = \mathrm{E}[X|\mathcal{F}_t] + \operatorname*{ess\,sup}_{Q\in\mathcal{Q}_\Lambda} \mathrm{E}_Q[-X|\mathcal{F}_t].$$

Therefore, conditional CVaRD is a lower range dominated conditional deviation measure with a one-to-one correspondence with the conditional value at risk. The uniquely determined set $\mathcal{Q}_\Lambda$ is called the risk envelope of the lower semicontinuous conditional deviation measure $\mathrm{CVaRD}_\Lambda(X|\mathcal{F}_t)$.

Proof. It is known that on the set $\mathcal{Q}_\Lambda$, the conditional CVaR has the representation

$$\mathrm{CVaR}_\Lambda(X|\mathcal{F}_t) = \operatorname*{ess\,sup}_{Q\in\mathcal{Q}_\Lambda}\{\mathrm{E}_Q[-X|\mathcal{F}_t]\}.$$

Then it follows that

$$\begin{aligned}\mathrm{CVaRD}_\Lambda(X|\mathcal{F}_t) &= \mathrm{CVaR}_\Lambda(X - \mathrm{E}[X|\mathcal{F}_t]|\mathcal{F}_t)\\ &= \operatorname*{ess\,sup}_{Q\in\mathcal{Q}_\Lambda}\{\mathrm{E}_Q[-X + \mathrm{E}[X|\mathcal{F}_t]|\mathcal{F}_t]\}\\ &= \mathrm{E}[X|\mathcal{F}_t] + \operatorname*{ess\,sup}_{Q\in\mathcal{Q}_\Lambda}\mathrm{E}_Q[-X|\mathcal{F}_t].\end{aligned}$$

Note that $\mathrm{E}_Q[\mathrm{E}[X|\mathcal{F}_t]|\mathcal{F}_t] = \mathrm{E}[X|\mathcal{F}_t]$ only holds if and only if $Q \in \mathcal{Q}_t$. Obviously, $\mathrm{CVaRD}_\Lambda(X|\mathcal{F}_t)$ is lower range dominated since $\dfrac{dQ}{dP} \geq 0$. Moreover, $\mathcal{Q}_\Lambda$ is a nonempty, closed and convex set, therefore, it is the risk envelop of $\mathrm{CVaRD}_\Lambda(X|\mathcal{F}_t)$.

□

To illustrate the representation, we provide some examples of calculation.

3.3.2 Calculation of CVaR and CVaRD

For some continuous distributions, it is possible to explicitly calculate the CVaR and CVaRD by definition.

1. **Normal distribution**. Suppose that X follows a normal distribution $N(\mu, \sigma^2)$. The quantile function of X is $F^{-1}(\alpha) = \mu + \sigma\Phi^{-1}(\alpha)$, where $\Phi^{-1}(\alpha)$ is

the quantile function of the standard normal distribution. Hence,

$$\begin{aligned}
\mathrm{CVaR}_\alpha(X) &= -\frac{1}{\alpha}\int_{-\infty}^{F^{-1}(\alpha)} \frac{x}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx \\
&= -\mu + \frac{\sigma}{\sqrt{2\pi}\alpha} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \Bigg|_{-\infty}^{F^{-1}(\alpha)} \\
&= -\mu + \frac{\sigma}{\sqrt{2\pi}\alpha} e^{-\frac{1}{2}(\Phi^{-1}(\alpha))^2},
\end{aligned}$$

and

$$\mathrm{CVaRD}_\alpha(X) = \mathrm{E}[X] + \mathrm{CVaR}_\alpha(X) = \frac{\sigma}{\sqrt{2\pi}\alpha} e^{-\frac{1}{2}(\Phi^{-1}(\alpha))^2}.$$

2. **Lognormal distribution**. Suppose that X follows a log-normal distribution $lnN(\mu, \sigma^2)$. Then mean is $\mathrm{E}[X] = e^{\mu + \frac{1}{2}\sigma^2}$ and the quantile function is $F^{-1}(\alpha) = e^{\mu + \sigma\Phi^{-1}(\alpha)}$. Using the result that

$$\int_k^\infty x f(x) dx = e^{\mu + \frac{1}{2}\sigma^2} \Phi\left(\frac{\mu + \sigma^2 - \ln k}{\sigma}\right),$$

we have

$$\begin{aligned}
\mathrm{CVaR}_\alpha(X) &= -\frac{1}{\alpha}\int_0^{F^{-1}(\alpha)} x f(x) dx \\
&= -\frac{1}{\alpha}\left(e^{\mu + \frac{1}{2}\sigma^2} - e^{\mu + \frac{1}{2}\sigma^2} \Phi\left(\frac{\mu + \sigma^2 - \ln e^{\mu + \sigma\Phi^{-1}(\alpha)}}{\sigma}\right)\right) \\
&= -\frac{1}{\alpha} e^{\mu + \frac{1}{2}\sigma^2}\left(1 - \Phi\left(\frac{\sigma^2 - \sigma\Phi^{-1}(\alpha)}{\sigma}\right)\right) \\
&= -\frac{1}{\alpha} e^{\mu + \frac{1}{2}\sigma^2} \Phi\left(\Phi^{-1}(\alpha) - \sigma\right),
\end{aligned}$$

and

$$\mathrm{CVaRD}_\alpha(X) = \mathrm{E}[X] + \mathrm{CVaR}_\alpha(X) = e^{\mu + \frac{1}{2}\sigma^2}\left(1 - \frac{1}{\alpha}\Phi\left(\Phi^{-1}(\alpha) - \sigma\right)\right).$$

3. **Student's t distribution**. Suppose that X follows a symmetric Student's t distribution with ν degrees of freedom $t(\nu)$. The mean is zero and the quantile function is $F^{-1}(\alpha)$. Hence

$$\begin{aligned}\mathrm{CVaR}_\alpha(X) &= -\frac{1}{\alpha}\int_{-\infty}^{F^{-1}(\alpha)} \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\nu\pi}\Gamma\left(\frac{\nu}{2}\right)} x\left(1+\frac{x^2}{\nu}\right)^{-\frac{\nu+1}{2}} dx \\ &= -\frac{\nu}{2\alpha}\int_{-\infty}^{F^{-1}(\alpha)} \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\nu\pi}\Gamma\left(\frac{\nu}{2}\right)}\left(1+\frac{x^2}{\nu}\right)^{-\frac{\nu+1}{2}} d\left(1+\frac{x^2}{\nu}\right) \\ &= \frac{\sqrt{\nu}\Gamma\left(\frac{\nu+1}{2}\right)}{(\nu-1)\alpha\sqrt{\pi}\Gamma\left(\frac{\nu}{2}\right)}\left(1+\frac{(F^{-1}(\alpha))^2}{\nu}\right)^{\frac{1-\nu}{2}},\end{aligned}$$

and

$$\mathrm{CVaRD}_\alpha(X) = \frac{\sqrt{\nu}\Gamma\left(\frac{\nu+1}{2}\right)}{(\nu-1)\alpha\sqrt{\pi}\Gamma\left(\frac{\nu}{2}\right)}\left(1+\frac{(F^{-1}(\alpha))^2}{\nu}\right)^{\frac{1-\nu}{2}}.$$

4. **Exponential distribution**. Suppose that X follows an exponential distribution $Exp(\lambda)$. The mean is $1/\lambda$ and the quantile function is $F^{-1}(\alpha) = \dfrac{-\ln(1-\alpha)}{\lambda}$. Hence

$$\begin{aligned}\mathrm{CVaR}_\alpha(X) &= -\frac{1}{\alpha}\int_0^{F^{-1}(\alpha)} \lambda x e^{-\lambda x} dx \\ &= F^{-1}(\alpha)e^{-\lambda F^{-1}(\alpha)} + \frac{1}{\lambda}e^{-\lambda F^{-1}(\alpha)} \\ &= \frac{1-\ln(1-\alpha)}{\lambda(1-\lambda)},\end{aligned}$$

and

$$\mathrm{CVaRD}_\alpha(X) = 1/\lambda + \frac{1-\ln(1-\alpha)}{\lambda(1-\lambda)}.$$

5. **Gamma distribution**. Suppose that X follows the gamma distribution $G(k,\theta)$. The mean is $k\theta$ and the quantile function is $F^{-1}(\alpha)$. Using the result

that

$$\int_0^x t^{s-1}e^{-t}dt = x^s\Gamma(s)e^{-x}\sum_{k=0}^{\infty}\frac{x^k}{\Gamma(s+k+1)},$$

we have

$$\begin{aligned}\mathrm{CVaR}_\alpha(X) &= -\frac{1}{\alpha}\int_0^{F^{-1}(\alpha)}\frac{1}{\Gamma(k)\theta^k}x^k e^{-x/\theta}dx\\ &= -\frac{1}{\alpha\Gamma(k)\theta^{k-1}}\int_0^{\frac{F^{-1}(\alpha)}{\theta}}x^k e^{-t}dt\\ &= -\frac{(k+1)(F^{-1}(\alpha))^{k+1}e^{-F^{-1}(\alpha)/\theta}}{\alpha\theta^{2k}}\sum_{k=0}^{\infty}\frac{(F^{-1}(\alpha))^k}{\Gamma(\alpha+k+2)}\end{aligned}$$

and

$$\mathrm{CVaRD}_\alpha(X) = k\theta - \frac{(k+1)(F^{-1}(\alpha))^{k+1}e^{-F^{-1}(\alpha)/\theta}}{\alpha\theta^{2k}}\sum_{k=0}^{\infty}\frac{(F^{-1}(\alpha))^k}{\Gamma(\alpha+k+2)\theta^k}.$$

3.3.3 Dynamic Deviation Measures

In this section we investigate conditional deviation measures in a dynamic framework where successive measurements are performed. Consider a finite set of time $\{t = 0, 1, \cdots, T\}$, we introduce a filtration $(\mathcal{F}_t)_{t=0,\cdots,T}$ where $\mathcal{F}_t$ models the information available at time t. Let $(L_t^\infty)^+$ be the set of all positive $\mathcal{F}_t$-measurable P-a.s. bounded random variables on which the essential supremum can be taken for each $0 \le t \le T$.

Definition 3.3.2. A dynamic deviation measure is a family $(\mathcal{D}_t)_{t=0,\cdots,T}$ such that $\mathcal{D}_t : L^2 \to (L_t^\infty)^+$ is a conditional deviation measure.

A dynamic deviation measure maps a random variable $X \in L^2$ into adapted process $(\mathcal{D}_t(X))_{t=0,\cdots,T}$ and can be seen as the result of the variability of risk assessment of a financial position through time. Definition (3.3.2) is quite general and it is interesting to look at some properties linking different components of a dynamic deviation measure such as the supermartingale property:

Supermartingale: For any $X \in L^2$ and $0 \leq t \leq T-1$,

$$\mathcal{D}_t(X) \geq \mathrm{E}[\mathcal{D}_{t+1}(X)|\mathcal{F}_t].$$

The semimartingale property can be interpreted as follows: as time evolves, the information about the financial postion X increases. This would lower the variability of the perceived riskiness - not almost surely, but in (conditional) mean.

There is no reference defining dynamic deviation measures and discussing their properties. We provide the definition in the next chapter and further discuss the dynamic deviation measures.

3.4 Summary

In this chapter, we have examined the extension of static risk and deviation measures, focusing on conditional and dynamic deviation measures. The motivation is to incorporate the dynamically available information into the process of assessing the riskiness of a financial position through time. Partially, we defined the conditional and dynamic deviation measures and provided a new representation result for the conditional CVaRD. Examples were provided to demonstrate its calculation.

Chapter 4

Risk Management for Income Streams

In this chapter, we provide an overview of the literature about dynamic risk and deviation measures in discrete time and introduce many new results. It has become obvious that risk measures have been thoroughly dealt with theoretically while there is only limited literature discussing deviation measures theoretically and practically. On the other hand, there is still a lack of practical dynamic risk and deviation measures to deal with processes which take into account the sequential development of incomes and payments over time. This is especially the case for dynamic deviation measures. To overcome this drawback, we generalize the concept of risk and the corresponding deviation measures developed by (Pflug and Ruszczyński (2005) and Pflug (2006)). In these works, a risk measure and its corresponding deviation measure for some financial models are introduced. Here we introduce the intuitive, dynamic counterparts of these risk and deviation measures. They are characterized by some optimization problems, which are solved by the method of Markov decision processes.

4.1 Dynamic Risk Measures for Income Process

In this part, we give a short overview on some important properties of dynamic risk measures in discrete and finite time, based on the stochastic process $Y = (Y_1, \cdots, Y_T)$ describing future incomes in periods $t = 1, \cdots, T$, defined on a probability space $(\Omega, \mathcal{F}, P)$. We introduce a filtration $(\mathcal{F}_t)_{t=0,1,\cdots,T}, T \in \mathbb{N}$, on $(\Omega, \mathcal{F}, P)$ where $\mathbb{N}$ is the set of all natural numbers and $\mathcal{F}_t$ models the information available at time t. Moreover, assume that $\mathcal{F}_0$ is trivial and $\mathcal{F}_T = \mathcal{F}$. A dynamic risk measure for an $(\mathcal{F}_t)_{t=0,1,\cdots,T}$-adapted process then is a mapping that assigns at every time point t an $\mathcal{F}_t$-measurable risk to the process, such that the sequence of the values of the risk measures is also an $(\mathcal{F}_t)_{t=0,1,\cdots,T}$-adapted process. Furthermore, define the space of all integrable processes in discrete time as

$$\chi := \{(Y_1, \cdots, Y_T) | Y_t \in L^1(\Omega, \mathcal{F}_t, P), \quad t = 1, \cdots, T\},$$

and define the space of all integrable and predictable processes starting at time $t+1$ for $t \in 0, 1, \cdots, T-1$ as

$$\chi^{T-t} := \{(X_t, \cdots, X_{T-1}) | X_k \in L^1(\Omega, \mathcal{F}_t, P), \quad k = t, \cdots, T-1\}.$$

A dynamic risk measure should satisfy the following two elementary properties.

Definition 4.1.1. Let $\rho : \Omega \times \{0, 1, \cdots, T-1\} \times \chi \to \mathbb{R} \cup \{\infty\}$ be a mapping and set $\rho_t(Y)(\omega) = \rho(\omega, t, Y)$ for all $(\omega, t, Y) \in \Omega \times \{0, 1, \cdots, T-1\} \times \chi$. Then

$(\rho_t(Y))_{t=0,\cdots,T-1}$ is called a dynamic risk measure if $\rho_t(Y)$ is $\mathcal{F}_t$-measurable for all $t = 0, \cdots, T-1$ and if the following two conditions are satisfied:

(IOP) $(\rho_t(Y))_{t=0,\cdots,T-1}$ is independent of the past, i.e., for every $t = 0, 1, \cdots, T-1$, and $Y \in \chi, \rho_t(Y)$ does not depend on $Y_1, \cdots, Y_{t-1}$. This condition also implies that for every $Y \in \chi$,

$$\rho_t(Y_1, \cdots, Y_T) = \rho_t(0, \cdots, 0, Y_t, \cdots, Y_T), \quad t = 0, \cdots, T-1.$$

(MON) $(\rho_t(Y))_{t=0,\cdots,T-1}$ is monotone, i.e., for all $Y^{(1)}, Y^{(2)} \in \chi$ with $Y^{(1)} \leq Y^{(2)}$ for all $t = 0, \cdots, T-1$, it holds

$$\rho_t(Y^{(1)}) \geq \rho_t(Y^{(2)})$$

In the following, we introduce the most important properties of dynamic risk measures such as translation invariance, coherence and consistency. We further assume that there exists a sequence of discount factors $\{\gamma_k\}_{k=0,1,\cdots,T}$ and let $\mathrm{E}_k \in \chi$ represent the income process that is one at a fixed point of time $k \in \{1, \cdots, T\}$ and zero otherwise.

The economic interpretation of the values of the risk measure is based on the translation invariance properties.

(TI1) For all $Y \in \chi, t \in \{0, 1, \cdots, T-1\}$ and $I \in L^1(\Omega, \mathcal{F}, P)$ it holds

$$\rho_t(Y + I\,\mathrm{E}_k) = \rho_t(I) - \frac{\gamma_k}{\gamma_t} I, \quad k = t, \cdots, T.$$

(TI2) Let $t \in \{0, 1, \cdots, T-1\}$. For every process $I = (0, \cdots, 0, I_t, \cdots, I_T) \in \chi$ such that $\sum_{k=t}^{T} \frac{\gamma_k}{\gamma_t} I_k$ is $\mathcal{F}_t$-measurable it holds for all $Y \in \chi$,

$$\rho_t(Y + I) = \rho_t(Y) - \sum_{k=t}^{T} \frac{\gamma_k}{\gamma_t} I_k.$$

We obtain for (TI1) or (TI2) that

$$\rho_t(Y + \rho_t(Y)) = \rho_t(Y) - \rho_t(Y) = 0$$

In a multiperiod setting, at every time t a process $I = (0, \cdots, 0, I_t, \cdots, I_T)$ should be added to the income process Y to eliminate the risk, as long as the discounted sum of the future values of I is known at time t. We will see later in this chapter that the assumption of I being a predictable process is also reasonable. Property (TI1) and (TI1) are inspired by Riedel (2004), Frittelli and Scandolo (2006) and Mundt (2007) in which they use $\gamma_k = (1 + r)^{-k}$, where $r > -1$ is a constant interest rate.

Analogously to the static case, dynamic risk measures can be classified by the properties of convexity and subadditivity. Moreover, under homogeneity, convexity and subadditivity are equivalent.

(CVX) A dynamic risk measure $(\rho_t(Y))_{t=0,\cdots,T-1}$ is called convex if for all $Y^{(1)}, Y^{(2)} \in \chi$, $t = \{0, \cdots, T-1\}$ and $\Lambda \in L^\infty(\Omega, \mathcal{F}_t, P)$ with $0 \leq \Lambda \leq 1$ it holds

$$\rho_t(\Lambda Y^{(1)} + (1 - \Lambda) Y^{(2)}) \leq \Lambda \rho_t(Y^{(1)}) + (1 - \Lambda) \rho_t(Y^{(2)}).$$

(SUB) A dynamic risk measure $(\rho_t(Y))_{t=0,\cdots,T-1}$ is called subadditive if for all $Y^{(1)}, Y^{(2)} \in \chi$, $t = \{0, \cdots, T-1\}$, it holds

$$\rho_t(Y^{(1)} + Y^{(2)}) \leq \rho_t(Y^{(1)}) + \rho_t(Y^{(2)}).$$

(HOM) A dynamic risk measure $(\rho_t(Y))_{t=0,\cdots,T-1}$ is called homogeneous if for all $Y \in \chi$, $t = \{0, \cdots, T-1\}$ and $\Lambda \in L^\infty(\Omega, \mathcal{F}_t, P)$ with $\Lambda \geq 1$, it holds

$$\rho_t(\Lambda Y) = \Lambda \rho_t(Y).$$

The following properties are essential tools for deriving representation and characterization theorems for dynamic risk measures.

(TCS) A dynamic risk measure $(\rho_t(Y))_{t=0,\cdots,T-1}$ is called time consistent if for all stopping times σ, τ on $\{0, 1, \cdots, T\}$ with $\sigma \leq \tau$, all processes $Y \in \chi$ and $I \in L^1(\Omega, \mathcal{F}_\tau, P)$, it holds

$$\rho_\sigma(Y + I\,\mathrm{E}_\tau) = \rho_\sigma\left(Y + \frac{\gamma_\tau}{\gamma_T} I\,\mathrm{E}_T\right)$$

(DCS) A dynamic risk measure $(\rho_t(Y))_{t=0,\cdots,T-1}$ is called dynamically consistent if for all processes $Y^{(1)}, Y^{(2)} \in \chi$ with $Y_t^{(1)} = Y_t^{(2)}$ for a $t \in \{0, 1, \cdots, T-1\}$, it holds

$$\rho_{t+1}(Y^{(1)}) = \rho_{t+1}(Y^{(2)}) \Rightarrow \rho_t(Y^{(1)}) = \rho_t(Y^{(2)}).$$

Definition 4.1.2. A dynamic risk measure $(\rho_t(Y))_{t=0,\cdots,T-1}$ is a dynamic coherent risk measure if it is translation invariant, subadditive and homogeneous.

4.2 Definition of the Dynamic Risk and Deviation Measures

Let $Y = (Y_1, \cdots, Y_T)$ be a stochastic process describing future incomes in periods $t = 1, \cdots, T$, defined on a probability space $(\Omega, \mathcal{F}, P)$. In addition, there exists an increasing sequence of sub-σ-algebras $(\mathcal{F}_t), t = 1, \cdots, T$, such that Y_t is $\mathcal{F}_t$-measurable. That is, the stochastic process $\{Y_t\}, t = 1, \cdots, T$, is adapted to the filtration $(\mathcal{F}_t), t = 1, \cdots, T$. At time $t-1$, a commitment is made to consume an amount a_{t-1} during the next period t. The consumption in period t is discounted by a factor of $c_t \geq 0$. If the current asset is less than the commitment, a shortfall occurs. The insurance cost for one unit of shortfall is $q_t \geq 0$. The shortfall is redeemed immediately and the shortfall cost is subtracted from the consumption before period t. Any surplus from period t is carried over and increases the wealth of the next period. The final wealth is discounted by a factor of $c_T \geq 0$. We make the following assumption about the sequences $\{c_t\}$ and $\{q_t\}$:

$$c_{t+1} \leq c_t \leq q_t, \quad t = 1, \cdots, T. \tag{4.1}$$

Assumption (4.1) makes sure that earlier consumption is preferable to later consumption, and any consumption is better than a surplus at the end.

If the wealth of the company at time t is denoted as W_t, the accumulated, non-negative wealth of the company can be defined recursively by

$$W_0 \equiv 0, W_t = (W_{t-1} + Y_t - a_{t-1})_+, \quad t = 1, \cdots, T,$$

and the shortfall at period $t, t = 1, \cdots, T$, is given by

$$M_t = (W_{t-1} + Y_t - a_{t-1})_-.$$

In every period $t \in \{0, 1, \cdots, T-1\}$, starting with the wealth W_t, the company's purpose is to maximize the discounted future utility

$$\frac{1}{c_t}\left\{\sum_{k=t+1}^{T}[c_k a_{k-1} - q_k M_k] + c_{T+1}W_T\right\} \tag{4.2}$$

for every sequence of integrable and predictable decisions $(a_t, \cdots, a_{T-1}) \in \chi^{(T-t)}$. This optimization problem motivates the definition of a dynamic risk measure as the negative supremum over all strategies (or decisions) $(a_t, \cdots, a_{T-1}) \in \chi^{T-t}$ of the expected utility, and the corresponding deviation measure as the infimum over all the strategies of the expected utility with full information minus the expected utility.

Definition 4.2.1. Let $W_t \in \mathcal{F}_t$ be the wealth at time $t \in \{0, 1, \cdots, T-1\}$. Then the dynamic risk measure is defined for Y as

$$\begin{aligned}\rho_t(Y) = -\frac{1}{c_t}\operatorname*{ess\,sup}_{(a_t,\cdots,a_{T-1})\in\chi^{T-t}}\Bigg\{ &- c_{t+1}W_t + \mathrm{E}\Bigg[\sum_{k=t+1}^{T}(c_k a_{k-1} - q_k M_k) \\ &+ c_{T+1}W_T|\mathcal{F}_t\Bigg]\Bigg\}.\end{aligned} \tag{4.3}$$

The corresponding dynamic deviation measure is defined for Y as

$$\begin{aligned}\mathcal{D}_t(Y) = \frac{1}{c_t}\operatorname*{ess\,inf}_{(a_t,\cdots,a_{T-1})\in\chi^{T-t}}\Bigg\{ &c_{t+1}W_t + \mathrm{E}\Bigg[\sum_{k=t+1}^{T}[c_k(Y_k - a_{k-1}) + q_k M_k] \\ &- c_{T+1}W_T|\mathcal{F}_t\Bigg]\Bigg\}.\end{aligned} \tag{4.4}$$

Proposition 4.2.1. For the dynamic risk measure $\rho_t(Y)$, we have the following:

(i) It always holds that

$$-\frac{1}{c_t}\sum_{k=t+1}^{T} c_k \,\mathrm{E}[W_k|\mathcal{F}_t] \le \rho_t(Y) \le -\frac{1}{c_t}\sum_{k=t+1}^{T} c_{k+1}\,\mathrm{E}[Y_k|\mathcal{F}_t].$$

(ii) $\mathrm{E}[|\rho_t(Y)|] < \infty$.

If we choose $c_k = q_k = c_{T+1}, \quad k = 1, \cdots, T$, then it holds that

$$\rho_t(Y) = -\sum_{k=t+1}^{T} c_{k+1}\,\mathrm{E}[Y_k|\mathcal{F}_t].$$

Proof. We only need to show part (i). The rest parts follow immediately from part (i) since we assume that $Y_k, k = 1, \cdots, T$, are integrable.

Because $c_{t+1} \le c_t \le q_t, t = 1, \cdots, T$, and $W_t - M_t = W_{t-1} + Y_t - a_{t-1}$ for an arbitrary strategy $(a_t, \cdots, a_{T-1}) \in \chi^{T-t}$, we have

$$\begin{aligned}
& -c_{t+1}W_t + \sum_{k=t+1}^{T}(c_k a_{k-1} - q_k M_k) + c_{T+1}W_T \\
\le\ & -c_{t+1}W_t + \sum_{k=t+1}^{T} c_k(a_{k-1} - M_k) + c_{T+1}W_T \\
=\ & -c_{t+1}W_t + \sum_{k=t+1}^{T} c_k(W_{k-1} + Y_k - W_k) + c_{T+1}W_T \\
=\ & -c_{t+1}W_t + \sum_{k=t+1}^{T} c_k Y_k + \sum_{k=t+1}^{T+1} c_k W_{k-1} - \sum_{k=t+1}^{T} c_k W_k \\
=\ & \sum_{k=t+1}^{T} c_k Y_k + \sum_{k=t+1}^{T} c_{k+1} W_k - \sum_{k=t+1}^{T} c_k W_k \\
\le\ & \sum_{k=t+1}^{T} c_k Y_k.
\end{aligned}$$

By this, the lower bound follows.

Now consider the admissible strategy $(a_t, \cdots, a_{T-1}) = (W_t, \cdots, W_{T-1}) \in \chi^{T-t}$. The wealth process becomes $W_k = Y_k, k = t+1, \cdots, T$, and $M_k = 0, k = t+1, \cdots, T$. The upper bound follows from

$$
\begin{aligned}
\rho_t(Y) &\leq -\frac{1}{c_t}\left\{-c_{t+1}W_t + \mathrm{E}\left[\sum_{k=t+1}^{T} c_k W_{k-1} + c_{T+1}W_T \middle| \mathcal{F}_t\right]\right\} \\
&= -\frac{1}{c_t}\,\mathrm{E}\left[\sum_{k=t+2}^{T} c_k Y_{k-1} + c_{T+1}Y_T \middle| \mathcal{F}_t\right] \\
&= -\frac{1}{c_t}\,\mathrm{E}\left[\sum_{k=t+1}^{T} c_{k+1} Y_k \middle| \mathcal{F}_t\right].
\end{aligned}
$$

□

Proposition 4.2.2. For the dynamic deviation measure $\mathcal{D}_t(Y)$, we have:

(i) It always holds that

$$\frac{1}{c_t}\sum_{k=t+1}^{T}(c_k - c_{k+1})\,\mathrm{E}[W_k|\mathcal{F}_t] \leq \mathcal{D}_t(Y) \leq \frac{1}{c_t}\sum_{k=t+1}^{T}(c_k - c_{k+1})\,\mathrm{E}[Y_k|\mathcal{F}_t].$$

(ii) $\mathcal{D}_t(Y)$ is nonnegative.

(iii) $\mathrm{E}[\mathcal{D}_t(Y)] < \infty$.

If we choose $c_k = q_k = c_{T+1}, k = 1, \cdots, T$, then it holds that

$$\mathcal{D}_t(Y) = 0.$$

Proof. We only need to show part (i). The rest parts follow immediately from part (i) since we assume that $Y_k, k = 1, \cdots, T$, are integrable.

Because $c_{t+1} \leq c_t \leq q_t, t = 1, \cdots, T$ and $W_t - M_t = W_{t-1} + Y_t - a_{t-1}$ for an arbitrary strategy $(a_t, \cdots, a_{T-1}) \in \chi^{T-t}$, we have

$$\begin{aligned}
& c_{t+1}W_t + \sum_{k=t+1}^{T} [c_k(Y_k - a_{k-1}) + q_k M_k] - c_{T+1}W_T \\
\geq\ & c_{t+1}W_t + \sum_{k=t+1}^{T} c_k(Y_k - a_{k-1} + M_k) - c_{T+1}W_T \\
=\ & c_{t+1}W_t + \sum_{k=t+1}^{T} c_k(W_k - W_{k-1}) - c_{T+1}W_T \\
=\ & c_{t+1}W_t + \sum_{k=t+1}^{T} c_k W_k - \sum_{k=t+1}^{T+1} c_k W_{k-1} \\
=\ & \sum_{k=t+1}^{T} (c_k - c_{k+1})W_k.
\end{aligned}$$

By this, the lower bound follows.

Now consider the admissible strategy $(a_t, \cdots, a_{T-1}) = (W_t, \cdots, W_{T-1}) \in \chi^{T-t}$. The wealth process becomes $W_k = Y_k, k = t+1, \cdots, T$, and $M_k = 0, k = t+1, \cdots, T$. The upper bound follows from

$$\begin{aligned}
\mathcal{D}_t(Y) &\leq \frac{1}{c_t}\left\{c_{t+1}W_t + \mathrm{E}\left[\sum_{k=t+1}^{T} c_k(Y_k - W_{k-1}) - c_{T+1}W_T \middle| \mathcal{F}_t\right]\right\} \\
&= \frac{1}{c_t}\mathrm{E}\left[\sum_{k=t+1}^{T} c_k Y_k - \sum_{k=t+2}^{T} c_k Y_{k-1} - c_{T+1}Y_T \middle| \mathcal{F}_t\right] \\
&= \frac{1}{c_t}\mathrm{E}\left[\sum_{k=t+1}^{T} c_k Y_k - \sum_{k=t+1}^{T} c_{k+1} Y_k \middle| \mathcal{F}_t\right] \\
&= \frac{1}{c_t}\sum_{k=t+1}^{T} (c_k - c_{k+1})\,\mathrm{E}\left[Y_k | \mathcal{F}_t\right].
\end{aligned}$$

$\square$

4.3 Properties of the Dynamic Risk and Deviation Measures

The dynamic risk measure (4.3) and deviation measure (4.4) at time t introduced in the previous section depend on the value of the wealth W_t. To simplify future arguments, we prove that it is possible to consider the optimization problem and the dynamic risk measure starting with $W_t = 0$.

Fix $t \in \{0, 1, \cdots, T-1\}$ and $(a_t, \cdots, a_{T-1}) \in \chi^{T-t}$. For the income process Y, we consider another sequence of decisions $(a'_t, \cdots, a'_{T-1})$ such that the accumulated wealth starting from t is recursively defined as

$$W'_t = 0, W'_{t+1} = (Y_{t+1} - a'_t)^+, W'_k = (W'_{k-1} + Y_k - a'_{k-1})^+, \quad k = t+2, \cdots .$$

By choosing $a'_t = a_t - W_t$ and $a'_k = a_k, k = t+1, \cdots, T-1$, we consequently have $W_{t+1} = (W_t + Y_{t+1} - a_t)^+ = W'_{t+1}$ and inductively $W_k = W'_k, k = t+1, \cdots, T$. Similarly, $M_k = M'_k, k = t+1, \cdots, T$.

Proposition 4.3.1. The dynamic risk measure (4.3) and the dynamic deviation measure (4.4) can be rewritten as

$$\rho_t(Y) = -\operatorname*{ess\,sup}_{(a_t,\cdots,a_{T-1})\in\chi^{T-t}} \frac{1}{c_t} \mathrm{E}\left[\sum_{k=t+1}^{T} (c_k a_{k-1} + q_k M_k) + c_{T+1} W_T \,\middle|\, \mathcal{F}_t\right] \tag{4.5}$$

and

$$\mathcal{D}_t(Y) = \operatorname*{ess\,inf}_{(a_t,\cdots,a_{T-1})\in\chi^{T-t}} \frac{1}{c_t} \mathrm{E}\left[\sum_{k=t+1}^{T} [c_k(Y_k - a_{k-1}) + q_k M_k] - c_{T+1} W_T \,\middle|\, \mathcal{F}_t\right] \tag{4.6}$$

respectively, where the wealth process W_t is defined as

$$W_{t+1} = (Y_{t+1} - a_t)^+, W_k = (W_{k-1} + Y_k - a_{k-1})^+, \quad k = t+2, \cdots, T.$$

Proof. We have

$$
\begin{aligned}
c_t\rho_t(Y) = {} & c_{t+1}W_t - \operatorname*{ess\,sup}_{(a_t,\cdots,a_{T-1})\in\chi^{T-t}} \mathrm{E}\left[\sum_{k=t+1}^{T}(c_k a_{k-1} - q_k M_k) + c_{T+1}W_T \,\middle|\, \mathcal{F}_t\right] \\
&= c_{t+1}W_t - \operatorname*{ess\,sup}_{(a_t,\cdots,a_{T-1})\in\chi^{T-t}} \mathrm{E}\Bigg[c_{t+1}a_t - q_{t+1}M_{t+1} \\
&\quad + \sum_{k=t+2}^{T}(c_k a_{k-1} - q_k M_k) + c_{T+1}W_T|\mathcal{F}_t\Bigg] \\
&= c_{t+1}W_t - \operatorname*{ess\,sup}_{(a_t,\cdots,a_{T-1})\in\chi^{T-t}} \mathrm{E}\Bigg[c_{t+1}(a'_t + W_t) - q_{t+1}M_{t+1} \\
&\quad + \sum_{k=t+2}^{T}(c_k a_{k-1} - q_k M_k) + c_{T+1}W_T|\mathcal{F}_t\Bigg] \\
&= -\operatorname*{ess\,sup}_{(a_t,\cdots,a_{T-1})\in\chi^{T-t}} \mathrm{E}\Bigg[c_{t+1}a'_t - q_{t+1}M_{t+1} \\
&\quad + \sum_{k=t+2}^{T}(c_k a_{k-1} + q_k M_k) + c_{T+1}W_T|\mathcal{F}_t\Bigg] \\
&= -\operatorname*{ess\,sup}_{(a'_t,\cdots,a'_{T-1})\in\chi^{T-t}} \mathrm{E}\left[\sum_{k=t+1}^{T}(c_k a'_{k-1} - q_k M'_k) + c_{T+1}W'_T \,\middle|\, \mathcal{F}_t\right] \qquad (4.7)
\end{aligned}
$$

Similarly,

$$
\begin{aligned}
c_t\mathcal{D}_t(Y) = {} & c_{t+1}W_t + \operatorname*{ess\,inf}_{(a_t,\cdots,a_{T-1})\in\chi^{T-t}} \mathrm{E}\left[\sum_{k=t+1}^{T}[c_k(Y_k - a_{k-1}) + q_k M_k] - c_{T+1}W_T \,\middle|\, \mathcal{F}_t\right] \\
&= c_{t+1}W_t + \operatorname*{ess\,inf}_{(a_t,\cdots,a_{T-1})\in\chi^{T-t}} \mathrm{E}\Bigg[c_{t+1}(Y_{t+1} - a_t) + q_{t+1}M_{t+1} \\
&\quad + \sum_{k=t+2}^{T}[c_k(Y_k - a_{k-1}) + q_k M_k] - c_{T+1}W_T|\mathcal{F}_t\Bigg] \\
&= c_{t+1}W_t + \operatorname*{ess\,inf}_{(a_t,\cdots,a_{T-1})\in\chi^{T-t}} \mathrm{E}\Bigg[c_{t+1}(Y_{t+1} - a'_t - W_t) + q_{t+1}M_{t+1} \\
&\quad + \sum_{k=t+2}^{T}[c_k(Y_k - a_{k-1}) + q_k M_k] - c_{T+1}W_T|\mathcal{F}_t\Bigg] \\
&= \operatorname*{ess\,inf}_{(a_t,\cdots,a_{T-1})\in\chi^{T-t}} \mathrm{E}\Bigg[c_{t+1}(Y_{t+1} - a'_t) + q_{t+1}M_{t+1}
\end{aligned}
$$

$$+ \sum_{k=t+2}^{T} [c_k(Y_k - a_{k-1}) + q_k M_k] - c_{T+1} W_T | \mathcal{F}_t \Bigg]$$

$$= \operatorname*{ess\,inf}_{(a'_t,\cdots,a'_{T-1})\in\chi^{T-t}} \mathrm{E}\left[\sum_{k=t+1}^{T} [c_k(Y_k - a'_{k-1}) + q_k M'_k] - c_{T+1} W'_T \middle| \mathcal{F}_t \right] \tag{4.8}$$

The optimization problems in equations (4.7) and (4.8) start now with $W'_t = 0$. $\square$

Proposition 4.3.2. If the income process $\{Y_t\}$ is predictable, then

$$\rho_t(Y) = -\frac{1}{c_t} \sum_{k=t+1}^{T} c_{k+1}\, \mathrm{E}[Y_k|\mathcal{F}_t], \quad t = \{0, 1, \cdots, T-1\},$$

and

$$\mathcal{D}_t(Y) = 0, \quad t = \{0, 1, \cdots, T-1\}.$$

Moreover, an optimal strategy is given by $(a^*_t, \cdots, a^*_{T-1}) = (Y_{t+1}, \cdots, Y_T) \in \chi^{T-t}$.

Proof. Using the proposed admissible strategy, the resulting wealth process becomes $W^*_{t+1} = \cdots = W^*_T = 0$. $\square$

Now we easily derive two simple properties, namely homogeneity and translation invariance, of the dynamic risk and deviation measures.

Proposition 4.3.3. Let $t \in \{0, 1, \cdots, T-1\}$.

(i) Let $\Lambda \in L^\infty(\Omega, \mathcal{F}_t, P)$ with $\Lambda > 0$, then

$$\rho_t(\Lambda Y) = \Lambda \rho_t(Y) \text{ and } D_t(\Lambda Y) = \Lambda \mathcal{D}_t(Y).$$

(ii) Let $I = (0, \cdots, 0, I_{t+1}, \cdots, I_T) \in \chi^T$ be a predictable process, then

$$\rho_t(Y + I) = \rho_t(Y) - \frac{1}{c_t} \sum_{k=t+1}^{T} c_{k+1}\, \mathrm{E}[I_k|\mathcal{F}_t] \text{ and } D_t(Y + I) = \mathcal{D}_t(Y).$$

Proof. (i) Let $(a_t, \cdots, a_{T-1}) \in \chi^{T-t}$. The the wealth process for the income process ΛY becomes

$$W_t^\Lambda = \Lambda W_t, W_{t+1}^\Lambda = (W_t^\Lambda + \Lambda Y_{t+1} - a_t)^+ = \Lambda \left(W_t + Y_{t+1} - \frac{a_t}{\Lambda} \right)^+,$$

$$W_k = (W_{k-1}^\Lambda + \Lambda Y_k - a_{k-1})^+ = \Lambda \left(W_{k-1} + Y_k - \frac{a_{k-1}}{\Lambda} \right)^+, \quad k = t+2, \cdots, T.$$

By defining $(a'_t, \cdots, a'_{T-1}) = \Lambda^{-1}(a_t, \cdots, a_{T-1}) \in \chi^{T-t}$, we have a one-to-one correspondence between the two strategies and consequently obtain

$$\begin{aligned}
\rho_t(\Lambda Y) = - \operatorname*{ess\,sup}_{(a_t, \cdots, a_{T-1}) \in \chi^{T-t}} \frac{1}{c_t} \Bigg\{ & - c_{t+1} W_t^\Lambda \\
& + \mathrm{E} \left[\sum_{k=t+1}^{T} (c_k a_{k-1} - q_k M_k^\Lambda) + c_{T+1} W_T^\Lambda \middle| \mathcal{F}_t \right] \Bigg\} \\
= -\Lambda \operatorname*{ess\,sup}_{(a'_t, \cdots, a'_{T-1}) \in \chi^{T-t}} \frac{1}{c_t} \Bigg\{ & - c_{t+1} W_{t+1} \\
& + \mathrm{E} \left[\sum_{k=t+1}^{T} (c_k a'_{k-1} - q_k M_k) + c_{T+1} W_T \middle| \mathcal{F}_t \right] \Bigg\} \\
= \Lambda \rho_t(Y), &
\end{aligned}$$

and

$$\begin{aligned}
D_t(\Lambda Y) = \operatorname*{ess\,inf}_{(a_t, \cdots, a_{T-1}) \in \chi^{T-t}} \frac{1}{c_t} \Bigg\{ & c_{t+1} W_t^\Lambda \\
& + \mathrm{E} \left[\sum_{k=t+1}^{T} [c_k(\Lambda Y_k - a_{k-1}) + q_k M_k^\Lambda] - c_{T+1} W_T^\Lambda \middle| \mathcal{F}_t \right] \Bigg\} \\
= \Lambda \operatorname*{ess\,inf}_{(a'_t, \cdots, a'_{T-1}) \in \chi^{T-t}} \frac{1}{c_t} \Bigg\{ & c_{t+1} W_{t+1} \\
& + \mathrm{E} \left[\sum_{k=t+1}^{T} [c_k(Y_k - a'_{k-1}) + q_k M_k] - c_{T+1} W_T \middle| \mathcal{F}_t \right] \Bigg\} \\
= \Lambda \mathcal{D}_t(Y) &
\end{aligned}$$

(ii) Let again $(a_t, \cdots, a_{T-1}) \in \chi^{T-t}$. Then the resulting wealth process for the income process $Y + I$ becomes

$$W_t^I = W_t, W_{t+1}^I = (W_t + Y_{t+1} + I_{t+1} - a_t)^+ = (W_t + Y_{t+1} - (a_t - I_{t+1}))^+,$$

$$W_k^I = (W_{k-1}^I + Y_k + I_k - a_{k-1})^+ = (W_{k-1}^I + Y_k - (a_{k-1} - I_k))^+, k = t+2, \cdots, T.$$

By defining $(a'_t, \cdots, a'_{T-1}) = (a_t - I_{t+1}, \cdots, a_{T-1} - Y_T) \in \chi^{T-t}$, we obtain

$$\begin{aligned}
\rho_t(Y+I) &= - \operatorname*{ess\,sup}_{(a_t,\cdots,a_{T-1})\in\chi^{T-t}} \frac{1}{c_t}\Bigg\{ - c_{t+1}W_t + \mathrm{E}\Bigg[\sum_{k=t+1}^{T} (c_k a_{k-1} - q_k M_k^I) \\
&\qquad + c_{T+1} W_T^I | \mathcal{F}_t\Bigg]\Bigg\} \\
&= - \operatorname*{ess\,sup}_{(a'_t,\cdots,a'_{T-1})\in\chi^{T-t}} \frac{1}{c_t}\Bigg\{ - c_{t+1}W_t + \mathrm{E}\Bigg[\sum_{k=t+1}^{T} [c_k(I_k + a'_{k-1}) - q_k M_k] \\
&\qquad + c_{T+1} W_T | \mathcal{F}_t\Bigg]\Bigg\} \\
&= \rho_t(Y) - \frac{1}{c_t}\sum_{k=t+1}^{T} c_{k+1}\,\mathrm{E}[Y_k|\mathcal{F}_t],
\end{aligned}$$

and

$$\begin{aligned}
D_t(Y+I) &= \operatorname*{ess\,inf}_{(a_t,\cdots,a_{T-1})\in\chi^{T-t}} \frac{1}{c_t}\Bigg\{ c_{t+1}W_t + \mathrm{E}\Bigg[\sum_{k=t+1}^{T} [c_k(Y_k + I_k - a_{k-1}) + q_k M_k^I] \\
&\qquad - c_{T+1} W_T^I | \mathcal{F}_t\Bigg]\Bigg\} \\
&= \operatorname*{ess\,inf}_{(a'_t,\cdots,a'_{T-1})\in\chi^{T-t}} \frac{1}{c_t}\Bigg\{ c_{t+1}W_t + \mathrm{E}\Bigg[\sum_{k=t+1}^{T} [c_k(Y_k - a'_{k-1}) + q_k M_k] \\
&\qquad - c_{T+1} W_T | \mathcal{F}_t\Bigg]\Bigg\} \\
&= \mathcal{D}_t(Y).
\end{aligned}$$

$\square$

4.4 Optimal Solution via Markov Decision Processes

In this section, we solve the optimization problem and obtain closed forms of the dynamic risk measure $\rho_t(Y)$ and the dynamic deviation measure $\mathcal{D}_t(Y)$. For the general income process, the solutions are obtained in Pflug and Ruszczyński (2005) and Pflug (2006) for both risk measure and deviation measure where the dual optimization method is applied. However these solutions are only for $t = 0$. We extend the results to $t > 0$, but restrict the income processes to depend on some underlying Markov chain.

Recall that for $t \in \{0, 1, \cdots, T\}$ and $Y \in \chi$,

$$\rho_t(Y) = -\frac{1}{c_t} \operatorname*{ess\,sup}_{(a_t,\cdots,a_{T-1})\in\chi^{T-t}} \Bigg\{ - c_{t+1}W_t + \mathrm{E}\Bigg[\sum_{k=t+1}^{T} (c_k a_{k-1} - q_k M_k) + c_{T+1}W_T | \mathcal{F}_t \Bigg] \Bigg\}$$

and

$$\mathcal{D}_t(Y) = \frac{1}{c_t} \operatorname*{ess\,inf}_{(a_t,\cdots,a_{T-1})\in\chi^{T-t}} \Bigg\{ c_{t+1}W_t + \mathrm{E}\Bigg[\sum_{k=t+1}^{T} [c_k(Y_k - a_{k-1}) + q_k M_k] - c_{T+1}W_T | \mathcal{F}_t \Bigg] \Bigg\},$$

where $\{W_t\}$ and $\{M_t\}$ are $\{\mathcal{F}_t\}$-adapted, $t = 0, 1, \cdots, T$.

To solve our Markov decision problem, we further require that the $\mathcal{F}_t$-adapted income process $\{Y_t\}$ has the Markov property

$$P(Y_t = y_t | \mathcal{F}_{t-1}) = P(Y_t = y_t | Y_{t-1} = y_{t-1}).$$

Also, we introduce the conditional value-at-risk risk measure CVaR_β as

$$\mathrm{CVaR}_\beta(Y) = \inf_{x\in\mathbb{R}} \left\{ \mathrm{E}\left[x + \frac{1}{\beta}[Y + x]_- \right] \right\},$$

where the minimum is taken over $x = \mathrm{VaR}_\beta(Y)$. The conditional value-at-risk deviation measure CVaRD_β is defined as

$$\mathrm{CVaRD}_\beta(Y) = \inf_{x\in\mathbb{R}} \left\{ \mathrm{E}\left[[Y - x]_+ + \frac{1-\beta}{\beta}[Y - x]_- \right] \right\}.$$

where the minimum is taken over $x = -\mathrm{VaR}_\beta(Y)$,

4.4.1 Solution for the Dynamic Risk Measure

Let us define all the quantities that are needed for a standard model of Markov decision processes and consider how to reformulate the dynamic deviation measure $\rho_t(Y)$.

- The state space is denoted by $S \subset \mathbb{R}^3$ and equipped with the σ-algebra $\mathcal{F}_S$. Let $s = (w, m, y) \in S$ be any element of the state space, where w, m, y represent realizations of the wealth process W_t, the shortfall process M_t and the income Markov chain Y_t respectively.

- The action space is $A \subset \mathbb{R}$ equipped with the σ-algebra $\mathcal{F}_A$, where $a \in A$ denotes the amount of consumption. There are no restrictions on the actions. For state $s \in S$, the set of admissible actions is $D(s) = A$ and the state-action space is $D = S \times A \subset \mathbb{R}^4$.

- The transition probability $P_t : D \times S \to [0,1], t = 1, \cdots, T$, is the conditional distribution of s_{t+1} given $(s_t, a_t) = (s, a) \in D$, formally defined by

$$P_t(s', s, a) = P(s_{t+1} = ((w+y'-a)^+, (w+y'-a)^-, y')|s_t = (w, m, y), a_t = a).$$

- The one-step reward function at time $t = 1, \cdots, T$, is a measurable mapping $r_t : D \to R$ defined by $r_t(s, a) = c_{t+1}a - q_t m, (s, a) \in D$.

- The terminal reward function is a measurable mapping $V_T : S \to \mathbb{R}$ given by

$$V_T(s) = c_{T+1}w - q_T m, \quad s = (w, m, y) \in S.$$

Furthermore, we introduce a set of admissible strategies.

Definition 4.4.1. For $t = 0, 1, \cdots, T-1$, the set of $(T-t)$-step admissible Markov policies is given by

$$F^{T-t} = \{\pi = (f_t, \cdots, f_{T-t})|f_k : S \to A \text{ is } (\mathcal{F}_S, \mathcal{F}_A)\text{−measurable}, \quad k = t, \cdots, T-1\}.$$

Because of the Markovian structure of all occurring random variables, the dynamic deviation measure $\rho_t(Y)$ becomes

$$
\begin{aligned}
&\rho_t(Y)\\
&= - \operatorname*{ess\,sup}_{(a_t,\cdots,a_{T-1})\in\chi^{T-t}} \frac{1}{c_t}\left\{-c_{t+1}W_t + E\left[\sum_{k=t+1}^{T}(c_k a_{k-1} - q_k M_k) + c_{T+1}W_T \middle| \mathcal{F}_t\right]\right\}\\
&= - \operatorname*{ess\,sup}_{(a_t,\cdots,a_{T-1})\in\chi^{T-t}} \frac{1}{c_t}\Bigg\{ - c_{t+1}W_t + c_{t+1}a_t + E\Bigg[\sum_{k=t+1}^{T-1}(c_{k+1}a_k - q_k M_k)\\
&\quad + c_{T+1}W_T - q_T M_T \Big| \mathcal{F}_t\Bigg]\Bigg\}\\
&= - \operatorname*{ess\,sup}_{(a_t,\cdots,a_{T-1})\in\chi^{T-t}} \frac{1}{c_t}\Bigg\{ - c_{t+1}W_t + q_t M_t + E\Bigg[\sum_{k=t}^{T-1}(c_{k+1}a_k - q_k M_k)\\
&\quad + c_{T+1}W_T - q_T M_T \Big| (W_t, M_t, Y_t) = (w,m,y)\Bigg]\Bigg\}\\
&= \frac{c_{t+1}}{c_t}w - \frac{q_t}{c_t}m - \operatorname*{ess\,sup}_{(a_t,\cdots,a_{T-1})\in\chi^{T-t}} \frac{1}{c_t}\left\{E\left[\sum_{k=t}^{T-1} r_k(X_k,a_k) + V_T(X_T)\middle| X_t = s\right]\right\}.
\end{aligned}
$$

We define the classic value function by

$$
V_{t,\pi}(s) = \mathrm{E}\left[\sum_{k=t}^{T-1} r_k(X_k,a_k) + V_T(X_T)\middle| X_t = s\right], \quad s \in S
$$

for every $\pi = (f_t, \cdots, f_{T-1}) \in F^{T-t}$ and

$$
V_t(s) = \sup_{\pi\in\chi^{T-t}} V_{t,\pi}(s), \quad s \in S.
$$

Then we obtain for $t = 0, 1, \cdots, T-1$,

$$
\rho_t(Y) = \frac{c_{t+1}}{c_t}W_t - \frac{q_t}{c_t}M_t - \frac{1}{c_t}V_t(W_t, M_t, Y_t). \tag{4.9}
$$

It is well known how to derive an explicit expression for the value functions. The following theorem is taken from Hernández-Lerma and Lasserre (1996) and is valid for general Borel spaces.

Theorem 4.4.1. Let $t \in \{0, 1, \cdots, T-1\}$ and $s \in S$ be fixed. If the functions $J_T(s') = V_T(s')$ and

$$J_k(s') = \sup_{a \in A}\{r_k(s', a) + \mathrm{E}[J_{k+1}(X_{k+1}|X_k = s', a_k = a)]\}, \quad s' \in S,\ k = t, \cdots, T-1,$$

are measurable and if the infimum is attained at $a^* = f_k^*(s')$, such that $f_k^* : S \to A$ is a $(\mathcal{F}_S, \mathcal{F}_A)$-measurable function, then $\pi^* = (f_t^*, \cdots, f_{T-1}^*) \in F^{T-t}$ is an optimal policy in the sense that

$$V_t(s) = V_{t,\pi^*}(s) = J_t(s)$$

Proof. This is Theorem 3.2.1 in Hernández-Lerma and Lasserre (1996) which is formulated and proven for $t = 0$ and the supremum problem. The general result holds, by either the proof given in Hernández-Lerma and Lasserre (1996) or some adoption to the argument for large t and the inferimum problem. □

Theorem 4.4.2. Let $t \in \{0, 1, \cdots, T\}$ and $(w, m, y) \in S$. The following results are true.

(i) The value function is given by

$$\begin{aligned} V_t(w, m, y) = c_{t+1}w - q_t m + \sum_{k=t+1}^{T} \mathrm{E}\Big[& c_{k+1}\,\mathrm{E}[Y_k|Y_{k-1}] \\ & - (c_k - c_{k+1})\,\mathrm{CVaR}_{\beta_k}(Y_k|Y_{k-1})|Y_t = y\Big]. \end{aligned}$$

(ii) The optimal policy $\pi^* = (f_t^*, \cdots, f_{T-1}^*)$ and the optimal Markov process $(X_t^*, \cdots, X_T^*)$ are given by

$$f_k^*(w', m', y') = w' - \mathrm{VaR}_{\beta_k}(Y_{k+1}|Y_k = y'), \quad s' = (w', m', y') \in S,$$

for $k = t, \cdots, T-1$ and the recursive relation

$$X_t^* = (w', m', y')$$

and for $k = t+1, \cdots, T$,

$$X_k^* = ((Y_k + \mathrm{VaR}_{\beta_k}(Y_k|Y_{k-1} = X_{k-1,3}^*))^+, (Y_k + \mathrm{VaR}_{\beta_k}(Y_k|Y_{t-1} = X_{k-1,3}^*))^-, Y_k).$$

Proof. The proof is by backward induction on t. The case t = T is trivial. We first consider the case when $t = T-1$, by

$$\begin{aligned}
&V_{T-1}(w, m, y) - c_T w + q_{T-1} m \\
&= \sup_{a \in A}\{-c_T w + q_{T-1} m + r_{T-1}((w, m, y), a) + \mathrm{E}[V_T(X_T) \\
&\quad |X_{T-1} = (w, m, y), a_{T-1} = a]\} \\
&= \sup_{a \in A}\{-c_T(w - a) + \mathrm{E}[c_{T+1}(w + y_T - a)^+ - q_T(w + y_T - a)^- \\
&\quad |X_{T-1} = (w, m, y), a_{T-1} = a]\} \\
&= \sup_{a \in A}\{\mathrm{E}[c_{T+1} y_T + (c_{T+1} - c_T)(w - a) + (c_{T+1} - q_T)(w + y_T - a)^- \\
&\quad |X_{T-1} = (w, m, y), a_{T-1} = a]\} \\
&= c_{T+1}\,\mathrm{E}[y_T|Y_{T-1} = y] - (c_T - c_{T+1}) \inf_{a \in A}\left\{E\left[(w - a) + \frac{q_T - c_{T+1}}{c_T - c_{T+1}}(w + y_T - a)^-\right.\right. \\
&\quad |X_{T-1} = (w, m, y), a_{T-1} = a]\} \\
&= c_{T+1}\,\mathrm{E}[y_T|Y_{T-1} = y] - (c_T - c_{T+1})\,\mathrm{CVaR}_{\beta_T}(Y_T|Y_{T-1} = y),
\end{aligned}$$

where

$$\beta_T = \frac{c_T - c_{T+1}}{q_T - c_{T+1}}.$$

The infimum is attained at

$$a^* = w - \mathrm{VaR}_{\beta_T}(Y_T|Y_{T-1} = y),$$

from which we derive that

$$X_T^* = ((Y_T + \mathrm{VaR}_{\beta_T}(Y_T|Y_{T-1} = y))^+, (Y_T + \mathrm{VaR}_{\beta_T}(Y_T|Y_{T-1} = y))^-, Y_T).$$

By lemma 1.1 in Mundt (2007), V_{T-1} and f_{T-1} are indeed measurable functions on S, so that Theorem 4.4.3 can be applied. Hence the assertion is true for $t = T - 1$.

Now assume that the assertion is true for $t \leq T - 1$. Together with the method of value iteration, this yields for $(w, m, y) \in S$,

$$\begin{aligned}
&V_{t-1}(w, m, y) - c_t w + q_{t-1} m \\
&= \sup_{a \in A}\{-c_t w + q_{t-1} m + r_{t-1}((w, m, y), a) + \mathrm{E}[V_t(X_t)|X_{t-1} = (w, m, y), a_{T-1} = a]\} \\
&= \sup_{a \in A}\{-c_t(w - a) + \mathrm{E}[c_{t+1}(w + Y_t - a)^+ - q_t(w + Y_t - a)^-|Y_{t-1} = y]\} \\
&\quad + \sum_{k=t+1}^{T} E[\mathrm{E}[c_{k+1}\,\mathrm{E}[Y_k|Y_{k-1}] - (c_k - c_{k+1})\,\mathrm{CVaR}_{\beta_k}(Y_k|Y_{k-1})|Y_t]|Y_{t-1} = y] \\
&= c_{t+1}\,\mathrm{E}[y_t|Y_{t-1} = y] - (c_t - c_{t+1})\,\mathrm{CVaR}_{\beta_t}(Y_t|Y_{t-1} = y) \\
&\quad + \sum_{k=t+1}^{T} E[c_{k+1}\,\mathrm{E}[Y_k|Y_{k-1}] - (c_k - c_{k+1})\,\mathrm{CVaR}_{\beta_k}(Y_k|Y_{k-1})|Y_{t-1} = y] \\
&= \sum_{k=t}^{T} \mathrm{E}[c_{k+1}\,\mathrm{E}[Y_k|Y_{k-1}] - (c_k - c_{k+1})\,\mathrm{CVaR}_{\beta_k}(Y_k|Y_{k-1})|Y_{t-1} = y],
\end{aligned}$$

where

$$\beta_k = \frac{c_k - c_{k+1}}{q_k - c_{k+1}}.$$

The infimum is attained at

$$a^* = w - \mathrm{VaR}_{\beta_t}(Y_t | Y_{t-1} = y),$$

and

$$X_t^* = ((Y_t + \mathrm{VaR}_{\beta_t}(Y_t|Y_{t-1} = y))^+, (Y_t + \mathrm{VaR}_{\beta_T}(Y_t|Y_{t-1} = y))^-, Y_t).$$

This completes the proof. □

The dynamic risk measure can now be represented as follows.

Corollary 4.4.1. For the dynamic risk measure we obtain for $t = 0, 1, \cdots, T-1$,

$$\rho_t(Y) = \frac{1}{c_t}\,\mathrm{E}\left[\sum_{k=t+1}^{T} \left(-c_{k+1}\,\mathrm{E}[Y_k|Y_{k-1}] + (c_k - c_{k+1})\,\mathrm{CVaR}_{\beta_k}(Y_k|Y_{k-1})\right) \middle| \, Y_t \right]. \tag{4.10}$$

Proof. This is a direct consequence of the previous theorem and (4.9). □

Remark. For $t = 0$, this result was also obtained in Pflug and Ruszczyński (2005), where a dual approach to the optimization problem is used. One advantage of our approach is the fact that we directly obtain the optimal values of the policy and the underlying wealth process. We are also able to derive a formula for $\rho_t(Y)$ when $t > 0$, where in Pflug and Ruszczyński (2005) only $\rho_0(Y)$ is investigated. One drawback of our method is that we have to assume a Markov structure for the income process Y_t.

4.4.2 Solution for the dynamic deviation measure

Let us define all the quantities that are needed for a standard model of Markov decision processes and consider how to reformulate the dynamic deviation measure $\mathcal{D}_t(Y)$.

- The state space is denoted by $S \subset \mathbb{R}^3$ and equipped with the σ-algebra $\mathcal{F}_S$. Let $s = (w, m, y) \in S$ be any element of the state space, where w, m, y represent realizations of the wealth process W_t, the shortfall process M_t and the income Markov chain Y_t respectively.

- The action space is $A \subset \mathbb{R}$ equipped with the σ-algebra $\mathcal{F}_A$, where $a \in A$ denotes the amount of consumption. There are no restrictions on the actions. For state $s \in S$, the set of admissible actions is $D(s) = A$ and the state-action space is $D = S \times A \subset \mathbb{R}^4$.

- The transition probability $P_t : D \times S \to [0, 1], t = 1, \cdots, T$ is the conditional distribution of s_{t+1} given $(s_t, a_t) = (s, a) \in D$, formally defined by

  $$P_t(s', s, a) = P(s_{t+1} = (w+y'-a)^+, (w+y'-a)^-, y')|s_t = (w, m, y), a_t = a).$$

- The one-step reward function at time $t = 1, \cdots, T$, is a measurable mapping $r_t : D \to \mathbb{R}$ defined by

  $$r_t(s, a) = c_t y - c_{t+1} a + q_t m, \quad (s, a) \in D.$$

- The terminal reward function is a measurable mapping $V_T : S \to \mathbb{R}$ given by

$$V_T(s) = c_T y - c_{T+1} w + q_T m, \quad s = (w, m, y) \in S.$$

Furthermore, we introduce a set of admissible strategies.

Definition 4.4.2. For $t = 0, 1, \cdots, T-1$, the set of $(T-t)$-step admissible Markov policies is given by

$$F^{T-t} = \{\pi = (f_t, \cdots, f_{T-t}) | f_k : S \to A \text{ is } (\mathcal{F}_S, \mathcal{F}_A)\text{−measurable}, \quad k = t, \cdots, T-1\}.$$

Because of the Markovian structure of all occurring random variables, the dynamic deviation measure $\mathcal{D}_t(Y)$ becomes

$$\begin{aligned}
\mathcal{D}_t(Y) = & \operatorname*{ess\,inf}_{(a_t,\cdots,a_{T-1})\in\chi^{T-t}} \frac{1}{c_t}\Bigg\{ c_{t+1}W_t + \mathrm{E}\Bigg[\sum_{k=t+1}^{T}[c_k(Y_k - a_{k-1}) + q_k M_k] \\
& - c_{T+1}W_T|\mathcal{F}_t\Bigg]\Bigg\} \\
= & \operatorname*{ess\,inf}_{(a_t,\cdots,a_{T-1})\in\chi^{T-t}} \frac{1}{c_t}\Bigg\{ c_{t+1}W_t - c_{t+1}a_t + \mathrm{E}\Bigg[\sum_{k=t+1}^{T-1}[c_k Y_k - c_{k+1}a_k + q_k M_k] \\
& + c_T y_T - c_{T+1}W_T + q_T M_T|\,\mathcal{F}_t\Bigg]\Bigg\} \\
= & \operatorname*{ess\,inf}_{(a_t,\cdots,a_{T-1})\in\chi^{T-t}} \frac{1}{c_t}\Bigg\{ c_{t+1}W_t - c_t Y_t - q_t M_t + \mathrm{E}\Bigg[\sum_{k=t}^{T-1}[c_k Y_k - c_{k+1}a_k + q_k M_k] \\
& + c_T y_T - c_{T+1}W_T + q_T M_T|\,(W_t, M_t, Y_t) = (w, m, z)\Bigg]\Bigg\} \\
= & \frac{c_{t+1}}{c_t}w - y - \frac{q_t}{c_t}m \\
& + \operatorname*{ess\,inf}_{(a_t,\cdots,a_{T-1})\in\chi^{T-t}} \frac{1}{c_t}\left\{\mathrm{E}\left[\sum_{k=t}^{T-1} r_k(X_k, a_k) + V_T(X_T)\middle|\, X_t = s\right]\right\}
\end{aligned}$$

By defining the classic value function by

$$V_{t,\pi}(s) = \mathrm{E}\left[\sum_{k=t}^{T-1} r_k(X_k, a_k) + V_T(X_T) \middle| X_t = s\right], \quad s \in S$$

for every $\pi = (f_t, \cdots, f_{T-1}) \in F^{T-t}$ and

$$V_t(s) = \inf_{\pi \in \chi^{T-t}} V_{t,\pi}(s), \quad s \in S,$$

we obtain for $t = 0, 1, \cdots, T-1$,

$$\mathcal{D}_t(Y) = \frac{c_{t+1}}{c_t} W_t - Y_t - \frac{q_t}{c_t} M_t + \frac{1}{c_t} V_t(W_t, M_t, Y_t). \tag{4.11}$$

It is well known how to derive an explicit expression for the value functions. The following theorem is taken from Hernández-Lerma and Lasserre (1996) and is valid for general Borel spaces.

Theorem 4.4.3. Let $t \in \{0, 1, \cdots, T-1\}$ and $s \in S$ be fixed. If the functions $J_T(s') = V_T(s')$ and

$$J_k(s') = \inf_{a \in A} \{r_k(s', a) + \mathrm{E}[J_{k+1}(X_{k+1} | X_k = s', a_k = a)]\}, \quad s' \in S,$$

defined for $k = t, \cdots, T-1$, are measurable and if the infimum is attained at $a^* = f_k^*(s')$, such that $f_k^* : S \to A$ is a $(\mathcal{S}, \mathcal{A})$-measurable function, then $\pi^* = (f_t^*, \cdots, f_{T-1}^*) \in F^{T-t}$ is an optimal policy in the sense that

$$V_t(s) = V_{t,\pi^*}(s) = J_t(s)$$

Proof. This is Theorem 3.2.1 in Hernándes-Lerma and Lasserre (1996) which is formulated and proven for $t = 0$ and the supreme problem. The general case holds by either the proof in Hernándes-Lerma and Lasserre (1996) or an adoption of the argument for large t and the infimum problem. □

Theorem 4.4.4. Let $t \in \{0, 1, \cdots, T\}$ and $(w, m, y) \in S$. The following results are true.

(i) The value function is given by

$$V_t(w, m, y) = c_t y - c_{t+1} w + q_t m + \sum_{k=t+1}^{T} (c_k - c_{k+1}) \,\mathrm{E}[\mathrm{CVaRD}_{\beta_k}(Y_k | Y_{k-1}) | Y_t = y].$$

(ii) The optimal policy $\pi^* = (f_t^*, \cdots, f_{T-1}^*)$ and the optimal Markov process $(X_t^*, \cdots, X_T^*)$ are given by

$$f_k^*(w', m', y') = w' - \mathrm{VaR}_{\beta_k}(Y_{k+1} | Y_k = y'), \quad s' = (w', m', y') \in S,$$

for $k = t, \cdots, T-1$, and the recursive relation

$$X_t^* = (w', m', y')$$

and for $k = t+1, \cdots, T$,

$$X_k^* = ((Y_k + \mathrm{VaR}_{\beta_k}(Y_k | Y_{k-1} = X_{k-1,3}^*))^+, (Y_k + \mathrm{VaR}_{\beta_k}(Y_k | Y_{t-1} = X_{k-1,3}^*))^-, Y_k).$$

Proof. The proof is by backward induction on t. The case t = T is trivial. We first consider the case when $t = T-1$, by

$$
\begin{aligned}
& V_{T-1}(w,m,y) + c_T w - c_{T-1} y - q_{T-1} m \\
&= \inf_{a\in A}\{c_T w - c_{T-1} y - q_{T-1} + r_{T-1}((w,m,y),a) + \mathrm{E}[V_T(X_T) \\
&\quad |X_{T-1} = (w,m,y), a_{T-1} = a]\} \\
&= \inf_{a\in A}\{c_T(w-a) + \mathrm{E}[c_T y_T - c_{T+1}(w+y_T-a)^+ + q_T(w+y_T-a)^- \\
&\quad |X_{T-1} = (w,m,y), a_{T-1} = a] \\
&= \inf_{a\in A}\{\mathrm{E}[c_T(w+y_T-a) - c_{T+1}(w+y_T-a)^+ + q_T(w+y_T-a)^- \\
&\quad |X_{T-1} = (w,m,y), a_{T-1} = a] \\
&= \inf_{a\in A}\{\mathrm{E}[c_T(w+y_T-a) + (c_T - c_{T+1})(w+y_T-a)^+ + (q_T - c_T)(w+y_T-a)^- \\
&\quad - c_T((w+y_T-a)^+ - c_T(w+y_T-a)^-)|X_{T-1} = (w,m,y), a_{T-1} = a] \\
&= \inf_{a\in A}\{\mathrm{E}[(c_T - c_{T+1})(w+y_T-a)^+ + (q_T - c_T)(w+y_T-a)^- \\
&\quad |X_{T-1} = (w,m,y), a_{T-1} = a] \\
&= (c_T - c_{T+1})\inf_{a\in A}\left\{E\left[[y_T-(a-w)]^+ + \frac{q_T - c_T}{c_T - c_{T+1}}[y_T-(a-w)]^- \middle| Y_{T-1} = y\right]\right\} \\
&= (c_T - c_{T+1})\,\mathrm{CVaRD}_{\beta_T}(Y_T|Y_{T-1} = y),
\end{aligned}
$$

where

$$\beta_T = \frac{c_T - c_{T+1}}{q_T - c_{T+1}}.$$

The infimum is attained at

$$a^* = w - \mathrm{VaR}_{\beta_T}(Y_T|Y_{T-1} = y),$$

from which we derive that

$$X_T^* = ((Y_T + \mathrm{VaR}_{\beta_T}(Y_T|Y_{T-1} = y))^+, (Y_T + \mathrm{VaR}_{\beta_T}(Y_T|Y_{T-1} = y))^-, Y_T).$$

By Lemma 1.1 in Mundt (2007), V_{T-1} and f_{T-1} are indeed measurable functions on S, so that Theorem 4.4.3 can be applied. Hence the assertion is true for $t = T - 1$.

Now assume that the assertion is true for $t \leq T - 1$. Together with the method of value iteration, this yields for $(w, m, y) \in S$,

$$\begin{aligned}
& V_{t-1}(w,m,y) + c_t w - c_{t-1} y - q_{t-1} m \\
= & \inf_{a \in A} \{ c_t w - c_{t-1} y - q_{t-1} m + r_{t-1}((w,m,y),a) + \mathrm{E}[V_t(X_t) \\
& |X_{t-1} = (w,m,y), a_{T-1} = a]\} \\
= & \inf_{a \in A} \{ c_t(w-a) + \mathrm{E}[c_t Y_t - c_{t+1}(w + Y_t - a)^+ + q_t(w + Y_t - a)^- | Y_{t-1} = y]\} \\
& + \sum_{k=t+1}^{T} (c_k - c_{k+1}) \, \mathrm{E}[\mathrm{E}[\mathrm{CVaRD}_{\beta_k}(Y_k|Y_{k-1})|Y_t]|Y_{t-1} = y] \\
= & (c_t - c_{t+1}) \, \mathrm{CVaRD}_{\beta_t}(Y_t|Y_{t-1} = y) \\
& + \sum_{k=t+1}^{T} (c_k - c_{k+1}) \, \mathrm{E}[\mathrm{CVaRD}_{\beta_k}(Y_k|Y_{k-1})|Y_{t-1} = y] \\
= & \sum_{k=t}^{T} (c_k - c_{k+1}) \, \mathrm{E}[\mathrm{CVaRD}_{\beta_k}(Y_k|Y_{k-1})|Y_{t-1} = y],
\end{aligned}$$

where

$$\beta_k = \frac{c_k - c_{k+1}}{q_k - c_{k+1}}.$$

The infimum is attained at

$$a^* = w - \mathrm{VaR}_{\beta_t}(Y_t|Y_{t-1} = y),$$

and

$$X_t^* = ((Y_t + \mathrm{VaR}_{\beta_t}(Y_t|Y_{t-1} = y))^+, (Y_t + \mathrm{VaR}_{\beta_T}(Y_t|Y_{t-1} = y))^-, Y_t).$$

This completes the proof □

The dynamic deviation measure can now be represented as follows.

Corollary 4.4.2. For the dynamic deviation measure, at $t = 0, 1, \cdots, T-1$, we have

$$\mathcal{D}_t(Y) = \frac{1}{c_t} \sum_{k=t+1}^{T} (c_k - c_{k+1})\, \mathrm{E}[\mathrm{CVaRD}_{\beta_k}(Y_k|Y_{k-1})|Y_t].$$

Proof. This is a direct consequence of the previous theorem and (4.11). □

Remark. For $t = 0$, this result was also obtained in Pflug (2006), where a dual approach to the optimization problem is used. One advantage of our approach is the fact that we directly obtain the optimal values of the policy and the underlying wealth process. Also we are able to derive a formula for $\mathcal{D}_t(Y)$ when $t > 0$, where in Pflug (2006) only $\mathcal{D}_0(Y)$ is investigated. One drawback of our method is that we have to assume a Markov structure of the income process Y_t.

4.4.3 Examples

Example 1. If Y_t follows a AR model $Y_t = a(Y_{t-1} - \mu) + \varepsilon_t$ where $\varepsilon_t \overset{iid}{\sim} N(0, \sigma^2)$, then the conditional distribution of Y_t given Y_{t-1} is $N(a(Y_{t-1} - \mu), \sigma^2)$. Hence at time t, the dynamic risk measure is given by

$$
\begin{aligned}
\rho_t(Y) &= \frac{1}{c_t}\,\mathrm{E}\left[\sum_{k=t+1}^{T} c_{k+1}a(Y_{k-1}-\mu)\right. \\
&\qquad \left.-(c_k-c_{k+1})\left[-a(Y_{k-1}-\mu)+\frac{\sigma}{\sqrt{2\pi}\beta_k}e^{-\frac{1}{2}(\Phi^{-1}(\beta_k))^2}\right]\,\middle|\, Y_t\right] \\
&= \frac{1}{c_t}\sum_{k=t+1}^{T} c_k a(E[Y_{k-1}|Y_t]-\mu) - \frac{1}{c_t}\sum_{k=t+1}^{T}(c_k-c_{k+1})\frac{\sigma}{\sqrt{2\pi}\beta_k}e^{-\frac{1}{2}(\Phi^{-1}(\beta_k))^2} \\
&= \frac{1}{c_t}\sum_{k=t+1}^{T} c_k a\left(a^{k-t-1}Y_t - \frac{(1-a^{k-t})\mu}{1-a} - \mu\right) \\
&\qquad - \frac{1}{c_t}\sum_{k=t+1}^{T}(c_k-c_{k+1})\frac{\sigma}{\sqrt{2\pi}\beta_k}e^{-\frac{1}{2}(\Phi^{-1}(\beta_k))^2} \\
&= \frac{1}{c_t}\sum_{k=t+1}^{T} c_k\left(a^{k-t}Y_t - \frac{a^2(1-a^{k-t-1})\mu}{1-a}\right) \\
&\qquad - \frac{1}{c_t}\sum_{k=t+1}^{T}(c_k-c_{k+1})\frac{\sigma}{\sqrt{2\pi}\beta_k}e^{-\frac{1}{2}(\Phi^{-1}(\beta_k))^2} \\
&= \frac{Y_t}{c_t}\sum_{k=t+1}^{T} c_k a^{k-t} - \frac{a^2\mu}{(1-a)c_t}\sum_{k=t+1}^{T} c_k(1-a^{k-t-1}) \\
&\qquad - \frac{1}{c_t}\sum_{k=t+1}^{T}(c_k-c_{k+1})\frac{\sigma}{\sqrt{2\pi}\beta_k}e^{-\frac{1}{2}(\Phi^{-1}(\beta_k))^2},
\end{aligned}
$$

and the deviation measure is given by

$$
\mathcal{D}_t(Y) = \frac{1}{c_t}\sum_{k=t+1}^{T}(c_k-c_{k+1})\frac{\sigma}{\sqrt{2\pi}\beta_k}e^{-\frac{1}{2}(\Phi^{-1}(\beta_k))^2}.
$$

Example 2. Consider the log-normal mean reverting model

$$
\log Y_t = a(\log Y_{t-1}-\mu) + \sigma\varepsilon_t,
$$

where $\varepsilon_t \overset{iid}{\sim} N(0,1)$. The conditional distribution of Y_t given Y_{t-1} is $\log N(a(\log Y_{t-1} - \mu), \sigma^2)$. Hence at time t, the risk measure is given by

$$\begin{aligned}
\rho_t(Y) &= \frac{1}{c_t} \mathrm{E}\left[\sum_{k=t+1}^{T} c_{k+1} e^{a(\log Y_{k-1} - \mu) + \frac{1}{2}\sigma^2} \right. \\
&\qquad \left. -(c_k - c_{k+1}) \left[-\frac{1}{\beta_k} e^{a(\log Y_{k-1} - \mu) + \frac{1}{2}\sigma^2 \Phi(\Phi^{-1}(\beta_k) - \sigma)} \right] \middle| \, Y_t \right] \\
&= \frac{1}{c_t} \sum_{k=t+1}^{T} \left(c_{k+1} e^{\frac{1}{2}\sigma^2} + (q_k - c_{k+1}) e^{\frac{1}{2}\sigma^2 \Phi(\Phi^{-1}(\beta_k) - \sigma)} \right) E\left[e^{a(\log Y_{k-1} - \mu)} \middle| \, Y_t \right] \\
&= \frac{1}{c_t} \sum_{k=t+1}^{T} \left(c_{k+1} e^{\frac{1}{2}\sigma^2} + (q_k - c_{k+1}) e^{\frac{1}{2}\sigma^2 \Phi(\Phi^{-1}(\beta_k) - \sigma)} \right) \\
&\qquad \times \exp\left\{ a^{k-t} \log y_t - \frac{\mu a \left(1 - a^{k-t}\right)}{1 - a} + \frac{\sigma^2 a^2 \left(1 - a^{2(k-t-1)}\right)}{2(1 - a^2)} \right\}
\end{aligned}$$

and the deviation measure is given by

$$\begin{aligned}
\mathcal{D}_t(Y) &= \frac{1}{c_t} \sum_{k=t+1}^{T} (c_k - c_{k+1}) E\left[e^{a(\log Y_{k-1} - \mu) + \frac{1}{2}\sigma^2} \left(1 - \frac{1}{\beta_k} \Phi(\Phi^{-1}(\beta_k) - \sigma) \right) \middle| \, Y_t \right] \\
&= \frac{1}{c_t} \sum_{k=t+1}^{T} (c_k - c_{k+1}) \left(1 - \frac{1}{\beta_k} \Phi(\Phi^{-1}(\beta_k) - \sigma) \right) E\left[e^{a(\log Y_{k-1} - \mu)} \middle| \, Y_t \right] \\
&= \frac{1}{c_t} \sum_{k=t+1}^{T} (c_k - c_{k+1}) \left(1 - \frac{1}{\beta_k} \Phi(\Phi^{-1}(\beta_k) - \sigma) \right) \\
&\qquad \times \exp\left\{ a^{k-t} \log y_t - \frac{\mu a \left(1 - a^{k-t}\right)}{1 - a} + \frac{\sigma^2 a^2 \left(1 - a^{2(k-t-1)}\right)}{2(1 - a^2)} \right\}.
\end{aligned}$$

4.5 A Stable Representation of the Dynamic Risk Measures and Deviation Measures

To state representation theorems, let P be the reference probability and $k \in \{1, \cdots, T\}$ be fixed. Further assume Q is a probability which is absolutely continuous with respect to P ($Q \ll P$) on $\mathcal{F}_k$ and $L_k^Q = \left.\dfrac{dQ}{dP}\right| \mathcal{F}_k$ is the resulting density. Then the density process of Q with respect to P is defined as

$$L_t^Q = \mathrm{E}[L_k^Q | \mathcal{F}_t], \quad t = 0, 1, \cdots, k.$$

Here L_t^Q is the density of Q with respect to P on $\mathcal{F}_t$ for every $t \in \{0, 1, \cdots, k\}$ and the process $(L_t^Q)_{t=0,\cdots,k}$ is an $(\mathcal{F}_t)_{t=0,\cdots,k}$-martingale.

Definition 4.5.1. A set of probability measures is called stable on $\mathcal{F}_k$ if for all Q_1, Q_2 which are equivalent to P with density processes $(L_t^{Q_1})_{t=0,\cdots,k}$ and $(L_t^{Q_2})_{t=0,\cdots,k}$ and all stopping time $\tau \leq k$ the process

$$L_t^{(\tau)} = \begin{cases} L_t^{Q_1} & t = 0, 1, \cdots, \tau, \\ L_\tau^{Q_1} \frac{L_t^{Q_2}}{L_\tau^{Q_2}} & t = \tau + 1, \cdots, k, \end{cases}$$

defines an element $Q^{(\tau)} \in \mathcal{Q}$ which is called the pasting of Q_1 and Q_2 in τ on $\mathcal{F}_k$

Definition 6.38 in (Föllmer and Schied, 2011) provides a way how we can directly obtain the probability $Q^{(\tau)}$ from Q_1 and Q_2 as

$$Q^{(\tau)}(A) = \mathrm{E}_{Q_1}[Q_2(A|\mathcal{F}_\tau)], \quad A \in \mathcal{F}_k.$$

Also for calculations it is useful to apply the result that for all $Q \ll P$ on $\mathcal{F}_k$ and stopping times ν, σ on $\{0, 1, \cdots, k\}$ with $\nu \geq \sigma$ it holds that

$$\mathrm{E}_Q[Z L_\sigma^Q | \mathcal{F}_\sigma] = \mathrm{E}[Z L_\nu^Q | \mathcal{F}_\sigma], \quad Z \in L^1(\Omega, \mathcal{F}_\nu, P).$$

This equality is hold in the almost surely sense with respect to P and hence with respect to Q.

For the dynamic risk measure

$$\rho_t(Y) = \frac{1}{c_t} \mathrm{E}\left[\sum_{k=t+1}^{T} \left(-c_{k+1} \mathrm{E}[Y_k|Y_{k-1}] + (c_k - c_{k+1}) \mathrm{CVaR}_{\beta_k}(Y_k|Y_{k-1})\right) \Big| Y_t \right],$$

let $\rho^{(k)}(X) = (1-\lambda_k)\,\mathrm{E}[-X] + \lambda_k \,\mathrm{CVaR}_{\beta_k}(X)$, where

$$\lambda_k = 1 - \frac{c_{k+1}}{c_k}, \beta_k = \frac{c_k - c_{k+1}}{q_k - c_{k+1}}.$$

Then $\rho^{(k)}(X)$ is a static law invariant coherent risk measure. Then we can rewrite our dynamic risk measure (4.10) as

$$\rho_t(Y) = \mathrm{E}\left[\sum_{k=t+1}^{T} \rho^{(k)}(Y_k|Y_{k-1}) \Big| Y_t \right]. \tag{4.12}$$

Let $\mathcal{P}$ be the set of all probability measures on Ω. Define the two sets $\mathcal{P}_k$ and $\mathcal{Q}_k$ of probability measures on each $\mathcal{F}_k$ for $k = \{1, \cdots, T\}$ as in Definition (3.1) and (3.2).

Lemma 4.5.1. Let $k = \{1, \cdots, T\}$,

$$\mathcal{Q}_k^* = \left\{ Q \in \mathcal{P}_k | \frac{c_{k+1}}{c_k} \leq L_k^Q \leq \frac{q_k}{c_k}, \mathrm{E}[L_k^Q|\mathcal{F}_{k-1}] = 1 \right\}. \tag{4.13}$$

(a) we have

$$\rho^{(k)}(X|\mathcal{F}_{k-1}) = \operatorname*{ess\,sup}_{Q \in \mathcal{Q}_k^*} \mathrm{E}_Q[-X|\mathcal{F}_{k-1}], \quad X \in L^1.$$

(b) For $t = 0, \cdots, k-1$, it holds

$$\mathrm{E}[\rho^{(k)}(X|\mathcal{F}_{k-1})|\mathcal{F}_t] = \operatorname*{ess\,sup}_{Q \in \mathcal{Q}_k^*} \mathrm{E}_Q[-X|\mathcal{F}_t], \quad X \in L^1.$$

Proof. Following the proof in example 6.2 in Ruszczyński and Shapiro (2006), we have for any $X \in L^1(\Omega, \mathcal{F}, P)$,

$$\begin{aligned}
\rho^{(k)}(X|\mathcal{F}_{k-1}) &= (1-\lambda_k)\,\mathrm{E}[-X|\mathcal{F}_{k-1}] + \lambda_k\,\mathrm{CVaR}_{\beta_k}(X|\mathcal{F}_{k-1}) \\
&= (1-\lambda_k)\,\mathrm{E}[-X|\mathcal{F}_{k-1}] + \lambda_k \operatorname*{ess\,inf}_{Z\in\mathcal{F}_{k-1}} \mathrm{E}\left\{Z + \frac{1}{\beta_k}[-X-Z]^+|\mathcal{F}_{k-1}\right\} \\
&= \operatorname{ess\,sup}\left\{\mathrm{E}[-LX|\mathcal{F}_{k-1}], L\in\mathcal{F}_k, \frac{c_{k+1}}{c_k} \le L \le \frac{q_k}{c_k}, \mathrm{E}[L|\mathcal{F}_{k-1}] = 1\right\} \\
&= \operatorname*{ess\,sup}_{Q\in\mathcal{Q}_k^*} \mathrm{E}_Q[-X|\mathcal{F}_{k-1}]
\end{aligned}$$

where $\mathcal{Q}_k^*$ is given by (4.13).

Now for any $B \in \mathcal{F}_{k-1} \subseteq \mathcal{F}_k$, Since

$$\mathrm{E}[L_k^Q|\mathcal{F}_{k-1}] = 1 \Leftrightarrow Q(B) = P(B), \quad \forall B \in \mathcal{F}_{k-1},$$

we have

$$Q = P \text{ on } \mathcal{F}_{k-1}.$$

Part (b) of this lemma easily follows from the result that if Q is equivalent to P on $\mathcal{F}_{k-1}$, then $\mathrm{E}_P[\mathrm{E}_Q[X|\mathcal{F}_{k-1}]] = \mathrm{E}_P[X]$. □

Mundt (2007) gave another proof of the above lemma. He also proved that

$$\mathcal{Q}_k^* = \left\{Q \in \mathcal{P}_k | \frac{c_{k+1}}{c_k} P(B) \le Q(B) \le \frac{q_k}{c_k} P(B), B \in \mathcal{F}_k, Q = P \text{ on } \mathcal{F}_{k-1}\right\}$$

is stable.

Lemma (4.5.1) and equation (4.12) together immediately yield the following stable representation result for the dynamic risk measure $\rho_t(Y)$.

Theorem 4.5.1. For every $t \in \{0, 1, \cdots, T-1\}$ it holds

$$\rho_t(Y) = \sum_{k=t+1}^{T} \frac{c_k}{c_t} \operatorname*{ess\,sup}_{Q \in \mathcal{Q}_k^*} \mathrm{E}_Q[-Y_k | Y_t]$$

For $k = 1, \cdots, T$, the set $\mathcal{Q}_k^*$ is stable on $\mathcal{F}_k$.

Based on the above stable representation of $\rho_t(Y)$, it is easy to prove that our dynamic risk measure is translation invariant, monotone, homogeneous and subadditive. Therefore, $\rho_t(Y)$ is a dynamic coherent risk measure.

For the dynamic deviation measure

$$\mathcal{D}_t(Y) = \frac{1}{c_t} \sum_{k=t+1}^{T} (c_k - c_{k+1})\, \mathrm{E}[\mathrm{CVaRD}_{\beta_k}(Y_k | Y_{k-1}) | Y_t]. \tag{4.14}$$

let $\mathcal{D}^{(k)}(X) = \lambda_k \, \mathrm{CVaRD}_{\beta_k}(X)$, where

$$\lambda_k = 1 - \frac{c_{k+1}}{c_k}, \beta_k = \frac{c_k - c_{k+1}}{q_k - c_{k+1}}.$$

Lemma 4.5.2. Let $k = \{1, \cdots, T\}$, then for

$$\mathcal{Q}_k^* = \left\{ Q \in \mathcal{P}_k | \frac{c_{k+1}}{c_k} \leq L_k^Q \leq \frac{q_k}{c_k}, \mathrm{E}[L_k^Q | \mathcal{F}_{k-1}] = 1 \right\},$$

(a) we have

$$\mathcal{D}^{(k)}(X | \mathcal{F}_{k-1}) = \mathrm{E}[X | \mathcal{F}_{k-1}] + \operatorname*{ess\,sup}_{Q \in \mathcal{Q}_k^*} \mathrm{E}_Q[-X | \mathcal{F}_{k-1}], \quad X \in L^1(\Omega, \mathcal{F}, P).$$

(b) For $t = 0, \cdots, k-1$, it holds

$$\mathrm{E}[\rho^{(k)}(X | \mathcal{F}_{k-1}) | \mathcal{F}_t] = \mathrm{E}[X | \mathcal{F}_t] + \operatorname*{ess\,sup}_{Q \in \mathcal{Q}_k^*} \mathrm{E}_Q[-X | \mathcal{F}_t], \quad X \in L^1(\Omega, \mathcal{F}, P).$$

Proof. We know that

$$\mathcal{D}^{(k)}(X|\mathcal{F}_{k-1}) = \lambda_k \,\mathrm{CVaRD}_{\beta_k}(X|\mathcal{F}_{k-1})$$
$$\rho^{(k)}(X|\mathcal{F}_{k-1}) = (1-\lambda_k)\,\mathrm{E}[-X|\mathcal{F}_{k-1}] + \lambda_k \,\mathrm{CVaR}_{\beta_k}(X|\mathcal{F}_{k-1})$$

By Theorem 3.3.1 and Lemma 4.5.1 (part (a)), it is easily to obtain that

$$\mathcal{D}^{(k)}(X|\mathcal{F}_{k-1}) = \mathrm{E}[X|\mathcal{F}_{k-1}] + \operatorname*{ess\,sup}_{Q\in\mathcal{Q}_k^*} \mathrm{E}_Q[-X|\mathcal{F}_{k-1}], \quad X \in L^1(\Omega,\mathcal{F},P),$$

where $\mathcal{Q}_k^*$ is given by (4.13). Now by Lemma 4.5.1 (part (b)),

$$\begin{aligned}\mathrm{E}[\mathcal{D}^{(k)}(X|\mathcal{F}_{k-1})|\mathcal{F}_t] &= \mathrm{E}[\mathrm{E}[X|\mathcal{F}_k]|\mathcal{F}_t] + \operatorname*{ess\,sup}_{Q\in\mathcal{Q}_k^*} \mathrm{E}[\mathrm{E}_Q[-X|\mathcal{F}_k]|\mathcal{F}_t] \\ &= \mathrm{E}[X|\mathcal{F}_t] + \operatorname*{ess\,sup}_{Q\in\mathcal{Q}_k^*} \mathrm{E}_Q[-X|\mathcal{F}_t].\end{aligned}$$

□

Since $\mathcal{D}^{(k)}(X|\mathcal{F}_{k-1})$ has a one-to-one correspondence with the conditional coherent risk measure $\rho^{(k)}(X|\mathcal{F}_{k-1})$, it is a lower range dominated conditional deviation measure. Now Lemma (4.5.2) and equation (4.14) immediately yield the following stable representation result for the dynamic deviation measure $\mathcal{D}_t(Y)$.

Theorem 4.5.2. For every $t \in \{0, 1, \cdots, T-1\}$ it holds

$$\mathcal{D}_t(Y) = \sum_{k=t+1}^{T} \frac{c_k}{c_t}\{\mathrm{E}[Y_k|Y_t] + \operatorname*{ess\,sup}_{Q\in\mathcal{Q}_k^*} \mathrm{E}_Q[-Y_k|Y_t]\}$$

Based on the above stable representation of $\mathcal{D}_t(Y)$, it is easy to prove that our dynamic deviation measure is a lower range dominated dynamic deviation measure.

4.6 Summary

This chapter is exclusively focused on characterizing and deriving new classes of dynamic risk and deviation measures for an income process. The tool used is the model of Markov decision processes. Both the dynamic risk and deviation measures are defined by means of the optimality criteria, which are solved recursively by the method of dynamic programming. Important properties of these dynamic risk and deviation measures are also discussed.

Chapter 5

Conclusion

5.1 Summary of Achievements

This PhD thesis is focused exclusively on risk management by means of Markov decision processes. Specifically, the research is focused on conditional risk and deviation measures, dynamic risk and deviation measures, and risk sensitive control of partially observable Markov decision processes.

The most fundamental notion in mathematical finance and risk management is that of risk. An associated, also important, notation is deviation. Although the goal of virtually every financial activity is to make a profit, such activity is always risky because the future financial state is always random and uncertain. Therefore, characterizing and quantifying risk and deviation is crucial for making rational and optimal financial decisions in the face of financial uncertainty.

The classic measures of risk and deviation have played important roles in mathematical finance and risk management in the past century, but they are single period and static. These two features put significant constraints on their applicability in practical situations involving uncertainty.

We have investigated significant extensions of single period, static risk and deviation measures along two important directions: the incorporation of information and the dynamic nature over time. When we represent information by a σ-algebra, we define conditional risk and deviation measures. For dynamic measures, we incorporate risk and deviation measures with the dynamic model of Markov decision processes to characterize and derive dynamic risk and deviation measures.

The stochastic model of Markov decision processes has made important contributions to mathematical finance and risk management. The classic model is risk neutral and the optimality criterion is additive. Since the model of Markov decision processes is very useful in risk management, expanding the classic model will inevitably broaden the horizon of its applications.

To apply to risk management, one significant extension is to incorporate the risk attitude into the model of Markov decision processes. Another significant contribution of this thesis is to extend the classic model of Markov decision processes simultaneously along two directions: partially observable Markov decision processes and risk sensitive Markov decision processes. Specifically, we have introduced and solved the model of risk sensitive partially observable Markov decision processes.

5.2 Future Research

In the literature of risk management, there is limited application of Markov decision processes. Since the model of Markov decision processes is powerful, further research of their applications in mathematical finance and risk management would be very fertile.

I am planning on some future research on the use of Markov decision processes in mathematical finance and risk management. One problem is to minimize the ruin probability by means of Markov decision processes. Here ruin will be characterized by some dynamic risk measures and the sequence of ruin probabilities will be characterized and derived by the model of Markov decision processes.

Another research problem is to further extend the dynamic risk and deviation measures introduced in this thesis. After applying risk sensitive Markov decision processes to characterize and derive dynamic risk measures, we hope to introduce totally new concepts of risk sensitive dynamic risk and deviation measures. The dynamic risk and deviation measures that are already derived in this thesis are multi-period and fully utilize the information available. The extension of them to the risk sensitive case would be more powerful by incorporating the decision maker's attitude towards risk. To the best of our knowledge, such as idea and the concept of risk sensitive dynamic risk measure does not exist in the current literature.

Bibliography

Acciaio, B., Föllmer, H. and Penner, I. (2010). Risk assessment for uncertain cash flows: Model ambiguity, discounting ambiguity. *Finance and Stochastics*, **16** 669–709.

Acciaio, B. and Penner, I. (2011). Dynamic risk measures. *Advanced Mathematical methods for Finance* 1–34.

Acerbi, C. (2002). Spectral measures of risk: a coherent representation of subjective risk aversion. *Journal of Banking & Finance*, **26** 1505–1518.

Agussurja, L. and Lau, H. C. (2011). A POMDP model for guiding taxi cruising in a congested urban city. In *the 10th Mexican international conference on Advances in Artificial Intelligence* (I. Batyrshin and G. Sidorov, eds.). MICAI2011, Puebla, Mexico, 415–428.

Arifoğlu, K. and Özekici, S. (2010). Optimal policies for inventory systems with finite capacity and partially observed Markov-modulated demand and supply processes. *European Journal of Operations Research*, **204** 421–438.

Artzner, P., Delbaen, F., Eber, J. M. and Heath, D. (1997). Thinking coherently. *Risk*, **10** 33–49.

Artzner, P., Delbaen, F., Eber, J. M. and Heath, D. (1999). Coherent measures of risk. *Mathematical Finance*, **9** 203–228.

Åström, K. J. (1965). Optimal control of Markov processes with incomplete state information. *Journal of the Mathematical Analysis and Applications*, **10** 174–205.

Bäuerle, N. and Riedel, U. (2011). *Markov Decision Processes with Applications to Finance*. Springer-Verlag, Heidelberg.

Bäuerle, N. and Riedel, U. (2014). More risk-sensitive Markov decision processes. *Mathematis of Operation Research*, **39** 105–120.

Bellman, R. E. (1957). *Dynamic Programming*. Princeton University Press, Princeton.

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

Brooks, A., Makarewko, A., Williams, S. and Durrant-Whyte, H. (2006). Parametric POMDPs for planning in continuous state spaces. *Robotics and Autonomous Systems*, **54** 887–897.

Carr, P., Geman, H. and Madan, D. (2001). Pricing and hedging in incomplete markets. *Journal of Financial Economics*, **62** 131–167.

Cerreia-Vioglio, S., Maccheroni, F., Marinacci, M. and Montrucchio, L. (2011). Uncertainty averse preferences. *Journal of Economic Theory*, **21** 743–774.

Cheridito, P., Delbaen, F. and Kupper, M. (2006). Dynamic monetary risk measures for bounded discrete-time processes. *Electronic Journal of Probability*, **11** 57–106.

Chong, E. K. P., Kreucher, C. M. and Hero III, A. O. (2009). Partially observable Markov decision process approximations for adaptive sensing. *Research Report, Institue National de Recherche en Informatique et en Automatique*, **19** 377–422.

Chung, K. J. and Sobel, M. J. (1987). Discounted MDP's distribution functions and exponential utility maximization. *SIAM Journal on Control and Optimization*, **25** 49–62.

Delbaen, F. (2002). Coherent measures of risk on general probability spaces. In *Advances in Finance and Stochastics. Essays in Honour of Dieter Sondermann.* Springer-Verlag, 1–37.

Detlefsen, K. and Scandolo, G. (2005). Conditional and dynamic convex risk measures. *Finance and Stochastics*, **9** 539–561.

Dynkin, E. A. and Yushkevich, A. A. (1979). *Controlled Markov Processes.* Springer, New York.

EI Karoui, N. and Ravaelli, C. (2009). Cash subadditive risk measures and interest rate ambiguity. *Mathematical Finance*, **19** 561–590.

Feinberg, E. A., Kasyanov, P. O. and Zgurovsky, M. Z. (2012). Optimality conditions for partially observable Markov decision processes. In *Continuous and Distributed Systems. Theory and Applications* (M. Z. Zgurovsky and V. A. Sadovnichiy, eds.), chap. 18. Springer, New York, 251–264.

Fernández-Gauchercurd, E., Arapostathis, A. and Marcus, S. I. (1991). On the average cost optimality equation and the structure of optimal policies for partially observable Markov decision processes. *Annals of Operations Research*, **29** 439–470.

Fishburn, P. C. (1970). *Utility Theory for Decision Making.* Wiley.

Föllmer, H. and Penner, I. (2006). Convex risk measures and the dynamics of their penalty functions. *Stochastics & Decisions*, **24** 61–96.

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

Föllmer, H. and Schied, A. (2008). Convex and coherent risk measures. Tech. rep., Humboldt University.

Föllmer, H. and Schied, A. (2011). *Stochastic Finance: An Introduction in Discreter Time (Third Edition)*. Walter de Gruyter, Berlin.

Franke, F., Härde, W. and G., S. (2000). *Measuring Risk in Complex Stochastic Systems*. Springer-Verlag, Berlin.

Frittelli, M. and Rosazza Gianin, E. (2002). Putting order in risk measures. *Journal of Banking & Finance*, **26** 1473–1486.

Frittelli, M. and Rosazza Gianin, E. (2004). *Dynamic convex risk measures*, chap. 12. New risk measures in the 21th century, Wiley Finance, 227–248.

Frittelli, M. and Scandolo, G. (2006). Risk measures and capital requirements for processes. *Mathematical Finance*, **16** 589–612.

Geman, H. (2002). Pure jump Lévy processes for asset modelling. *Journal of Banking & Finance*, **26** 1297–1316.

Goulionis, J. E. and Vozikis, A. (2012). Modeling decision making in clinical practice: a cost-effectiveness approach. *Revista Investigacion Operational*, **33** 29–37.

Hernández-Lerma, O. (1989). *Adaptive Markov Control Processes*. Springer, New York.

Hernández-Lerma, O. and Lasserre, J. B. (1996). *Discrete-Time Markovian Decision Processes: Basic Optimality Criteria*. Springer, Berlin.

HINDERER, K. (1970). *Foundations of Non-Stationary Dynamic Programming with Discrete Time Parameter*. Springer-Verlag, Berlin.

HOWARD, R. A. (1960). *Dynamic Programming and Markov Processes.* M.I.T Press, Cambridge, Massachusetts.

HOWARD, R. A. and MATHESONS, J. E. (1972). Risk-sensitive Markov decision processes. *Management Science*, **18** 356–369.

KLÖPPEL, S. and SCHWEIZER, M. (2007). Dynamic utility indifference valuation via convex risk measures. *Mathematical Finance*, **17** 599–627.

KREPS, D. M. (1977a). Decision problems with expected utility criteria, i: upper and lower convergent utility. *Mathematis of Operation Research*, **2** 45–53.

KREPS, D. M. (1977b). Decision problems with expected utility criteria, ii: stationarity. *Mathematis of Operation Research*, **2** 266–274.

KUPPER, M. and SCHACHERMAYER, W. (2009). Representation results for law invariance time consistent functions. *Mathematics and Financial Economics*, **2/3** 189–210.

LARSEN, K., PIRVU, T., SHREVE, S. and TUTUNCU, R. (2005). Satisfying convex risk limits by trading. *Finance and Stochastics*, **9** 177–195.

LOVEJOY, W. S. (1991). A survey of algorithm methods to partially observable markov decision processes. *Annals of Operation Research*, **28** 47–65.

MARKOWITZ, H. M. (1952). Portfolio selection. *Journal of Finance*, **7** 77–91.

MARKOWITZ, H. M. (1959). *Portfolio Selection*. Wiley, New York.

MERTON, R. C. (1989). *Continuous-Time Finance*. Basil Blackwell, Oxford.

MUNDT, A. P. (2007). *Dynamic Risk Management with Markov Decision Processes*. Ph.D. thesis, Universität Karlsruhe.

NAKAI, T. (2010). Sequential decision problem with partial maintenance on a partially observable Markov process. *Scientia Mathematicae Japonicae Online Online*, **283-292**.

OGRYCZAK, W. and RUSZCZYŃSKI, A. (1999). From stochastic dominance to mean-risk modes: semideviations as risk measures. *European Journal of Operations Research*, **116** 33–50.

OGRYCZAK, W. and RUSZCZYŃSKI, A. (2001). On consistency of stochastic dominance and mean semideviation models. *Mathematical Programming*, **89** 217–232.

OGRYCZAK, W. and RUSZCZYŃSKI, A. (2002). Dual stochastic dominance and related mean risk models. *SIAM Journal on Optimization*, **13** 60–78.

PARTHASARATHY, K. R. (1967). *Probability Measures on Metric Spaces*. Academic Press, New York.

Penner, I. (2007). *Dynamic convex risk measures: time consistency, prudence, and sustainability*. Ph.D. thesis, Humboldt-Universität zu Berlin.

Pflug, G. C. (2000). Some remarks on the value-at-risk and the conditional value-at-risk. In *Probabilistic Constrained Optimization: Methodology and Applications* (S. Uryasev, ed.). Kluwer Academic Publishers, 272–281.

Pflug, G. C. (2006). A value-of-information approach to measuring risk in multi-period economic activity. *Journal of Banking & Finance*, **30** 695–715.

Pflug, G. C. and Ruszczyński, A. (2005). Measuring risk for income streams. *Computational Optimization and Applications*, **32**.

Puterman, M. L. (2014). *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. John Wiley & Sons.

Rebonato, R. (2001). Managing model risk. In *Financial Times, Mastering Risk* (C. Alexander, ed.), vol. 2. Pearson Education, 82–116.

Rhenius, D. (1974). Incomplete information in Markovian decision models. *Annals of Statistics*, **2** 1327–1334.

Riedel, F. (2004). Dynamic coherent risk measures. *Stochastic Processes and their Applications*, **112** 185–200.

Ristic, B., Arulampalam, S. and Gorden, N. (2004). *Beyond the Kalman Filter: Particle Filters for Tracking Applications*. Artech House, Norwood.

Rockafellar, R. T. (2007). Coherent approaches to risk in optimization under uncertainty. In *Tutorials In Operations Research, INFORMS Tutorial*. Institute for Operations Research and the Management Sciences, Hanover, MD, 38–61.

Rockafellar, R. T. and Uryasev, S. (2002). Conditional value at risk for general loss distributions. *Journal of Banking & Finance*, **26** 1143–1171.

Rockafellar, R. T., Uryasev, S. and Zabarankin, M. (2006). Generalized deviations in risk analysis. *Finance and Stochastic*, **10** 51–74.

Roorda, B., Schumacher, J. M. and Engwerda, J. C. (2005). Coherent acceptability measures in multiperiod models. *Mathematical Finance*, **15** 589–612.

Roy, A. (1952). Safety first and the holding of assets. *Econometrica*, **20** 431–449.

Ruszczyński, A. (2010). Risk-averse dynamic programming for Markov decision processes. *Mathematical Programming*, **125** 235–261.

Ruszczyński, A. and Shapiro, A. (2006). Conditional risk mappings. *Mathematis of Operation Research*, **31** 544–561.

Sarykalin, S., Serraino, G. and Uryasev, S. (2008). VaR vs CVaR in risk management and optimization. In *Tutorials In Operations Research, INFORMS Tutorial*. Institute for Operations Research and the Management Sciences, Hanover, MD, 270–294.

Sawaragi, Y. and Yoshikawa, T. (1970). Discrete-time Markovian decision processes with incomplete state observations. *Annals of Mathematical Statistics*, **41** 78–86.

Smallwood, R. D. and Sondik, E. J. (1973). The optimal cotrol of partially observable Markov processes over a finite horizon. *Operation Research*, **21** 1071–1088.

Sornette, D. and Johansen, A. (2001). Significance of log-periodic precursors to financial crashes. *Quantitative Finance*, **1** 452–471.

Szegö, G. (2004). On the (non)acceptance of innovations. In *Risk measures for the 21st Century* (G. Szegö, ed.). Jonh Wiley & Sons, 1–10.

Tomberlin, D. (2010). Endangered seabird habitat management as a partially observable Markov decision process. *Marine Resource Economics*, **25** 93–104.

Tversky, A. and Simonson, I. (1993). Context-dependent preferences. *Management Science*, **39** 1179–1189.

Vozikis, A., Goulionis, J. E. and Benos, V. K. (2012). The partially observable Markov decision processes in healthcare: an application to patients with ischemic heart disease (IHD). *Operation Research - An International Journal*, **12** 3–14.

Weber, S. (2006). Distribution-invariant risk measures, information, and dynamic consistency. *Mathematical Finance*, **16** 419–441.

White, D. J. (1988). Mean, variance, and probabilistic criteria in finite Markov decision processes, a review. *Journal of Optimization Theory and Applications*, **56** 1–29.

White III, C. C. (1991). A survey of solution techniques for the partially observable Markov decision processes. *Annals of Operation Research*, **32** 215–230.

Yushkevich, A. A. (1976). Reduction of a controlled Markov model with incomplete data to a problem with complete information in the case of Borel state and control spaces. *Theory of Probability and Its Applications*, **21** 153–158.

Appendix A

List of symbols

- $(\Omega, \mathcal{F}, P)$: given probability space
- $L^k(\Omega, \mathcal{F}, P), k \in [1, \infty]$: space of real-valued random variables defined on $(\Omega, \mathcal{F}, P)$ such that $\|X\|_k < \infty$
- $L^\infty := L^\infty(\Omega, \mathcal{F}, P)$: space of essentially bounded random variables defined on $(\Omega, \mathcal{F}, P)$
- $L^1 := L^1(\Omega, \mathcal{F}, P)$: space of integrable random variables defined on $(\Omega, \mathcal{F}, P)$
- $L^2 := L^2(\Omega, \mathcal{F}, P)$: space of square integrable random variables defined on $(\Omega, \mathcal{F}, P)$
- $(\mathcal{F}_t)_{t=0,\cdots,T}$: a filtration on $(\Omega, \mathcal{F}, P)$ where $\mathcal{F}_t$ models the information available at time t
- L_t^∞: the set of all $\mathcal{F}_t$-measurable P-a.s. bounded random variables

- $(L_t^\infty)^+$: the set of all positive $\mathcal{F}_t$-measurable P-a.s. bounded random variables
- $\rho(X)$: risk measure of a financial state X
- $\mathcal{D}(X)$: deviation measure of a financial state X
- $\mathcal{Q}$: risk envelope
- $Q \sim P$: Q is equivalent to P
- $Q = P$: Q is equal to P.
- $X^+(X^-)$: positive (negative) part of a random variable X, i.e., $X^+ = \max(X, 0)$, $X^- = \max(-X, 0)$.

Appendix B

List of terms

- APT: arbitrage pricing theory
- CAPM: capital asset pricing model
- CVaR: conditional value at risk
- CVaRD: conditional value at risk deviation
- ES: expected shortfall
- MDP: Markov decision process
- P&L: profit and loss
- POMDP: partially observable Markov decision process
- RSMDP: risk sensitive Markov decision process
- RSPOMDP: risk sensitive partially observable Markov decision process

- SOSD: second order stochastic dominance
- SDF: stochastic discount factor
- TVaR: tail value at risk
- VaR: value at risk