FACULTY
OF MATHEMATICS
AND PHYSICS
Charles University

MASTER THESIS

Lucia Petrusova

Robust Methods in Portfolio Theory

Department of Probability and Mathematical Statistics

Supervisor of the master thesis:	RNDr. Martin Branda, Ph.D.
Study programme:	Mathematics
Study branch:	Probability, Mathematical Statistics and Econometrics

Prague 2016

I declare that I carried out this master thesis independently, and only with the cited sources, literature and other professional sources.

I understand that my work relates to the rights and obligations under the Act No. 121/2000 Sb., the Copyright Act, as amended, in particular the fact that the Charles University has the right to conclude a license agreement on the use of this work as a school work pursuant to Section 60 subsection 1 of the Copyright Act.

In .... date .... signature of the author

Title: Robust Methods in Portfolio Theory

Author: Lucia Petrusova

Department: Department of Probability and Mathematical Statistics

Supervisor: RNDr. Martin Branda, Ph.D., Department of Probability and Mathematical Statistics

Abstract: This thesis is concerned with the robust methods in portfolio theory. Different risk measures used in portfolio management are introduced and the corresponding robust portfolio optimization problems are formulated. The analytical solutions of the robust portfolio optimization problem with the lower partial moments (LPM), value-at-risk (VaR) or conditional value-at-risk (CVaR), as a risk measure, are presented. The application of the worst-case conditional value-at-risk (WCVaR) to robust portfolio management is proposed. This thesis considers WCVaR in the situation where only partial information on the underlying probability distribution is available. The minimization of WCVaR under mixture distribution uncertainty, box uncertainty, and ellipsoidal uncertainty are investigated. Several numerical examples based on real market data are presented to illustrate the proposed approaches and advantage of the robust formulation over the corresponding nominal approach.

Keywords: robust methods, portfolio selection, risk measures, conditional value-at-risk.

I would like to thank my supervisor RNDr. Martin Branda, Ph.D. for his ideas and guidance throughout the preparation of this thesis. I am also grateful to my family and friends for their patience and support.

Contents

Introduction **2**

1 Risk Measures **4**

1.1 Lower Partial Moments 5

1.1.1 Upper Bounds for the Univariate Cases 5

1.2 Value-at-risk 7

1.3 Conditional Value-at-risk 8

1.4 Worst-case VaR and Worst-case CVaR 8

2 Minimization of Risk Measures **10**

2.1 Minimization of CVaR 10

2.2 Minimization of Worst-case CVaR 11

2.2.1 Mixture Distribution 12

2.2.2 Discrete Distribution 16

2.2.3 Box Uncertainty in Discrete Distribution 16

2.2.4 Ellipsoidal Uncertainty in Discrete Distribution 18

3 Robust Portfolio Selection **20**

3.1 Portfolio Selection with CVaR 21

3.2 Robust Portfolio Selection Using LPM 21

3.2.1 Explicit Solution of the Robust Portfolio Problem 22

3.3 Robust Portfolio Selection Using WCVaR 25

3.3.1 Mixture Distribution Uncertainty 26

3.3.2 Box Uncertainty in Discrete Distribution 26

3.3.3 Ellipsoidal Uncertainty in Discrete Distribution 27

4 Numerical Application **29**

4.1 CVaR Approach 29

4.2 WCVaR under Box Uncertainty 32

4.3 WCVaR under Mixture Distribution 34

Conclusion **38**

Bibliography **39**

List of Figures **40**

List of Tables **41**

Attachments **42**

Introduction

On the financial markets, investors constantly face to a trade-off between adjusting potential returns for higher risk. Recently, there is a number of ways that risk can be defined and measured.

In 1952, Harry Markowitz introduced modern porfolio theory [13], or *mean-variance analysis*. Due to symmetrical nature of the variance, which is the reason why the variance does not differentiate the gain from the loss, even Markowitz himself later proposed using the semivariance instead. To be better convenient for different risk profiles of the investors, Bawa [2] and Fishburn [8] introduced a class of downside risk measures known as the lower partial moments, or LPM. Value-at-risk, or VaR, is a popular measure of risk in financial risk management. However, VaR has been critized in recent years in several aspects. VaR is not subadditive in general distribution case and thus it is not a coherent risk measure in the sense of Artzner [1]. A very serious shortcoming of VaR is that it is just a percentile of a loss distribution, so it does not show the nature of extreme losses exceeding it. These troubles motivated the search for a better measure of risk than VaR for practical applications. Conditional value-at-risk, or CVaR, roughly defined as the mean of the tail distribution exceeding VaR, is a measure of risk with significant advantages over VaR. It is able to quantify dangers beyond VaR and it is coherent. Rockafellar and Uryasev [15] introduced a fundamental minimization formula for CVaR and showed that CVaR can be calculated by minimizing a more tractable auxiliary function without predetermining the corresponding VaR. Moreover, VaR can be calculated as a by-product. The CVaR minimization formula usually results in convex programs, and even linear programs. Therefore, CVaR attracted much attention in recent years and is applied to financial optimization and risk management.

As Black and Litterman [3] noticed, in the classical mean variance model, the portfolio selection is very sensitive to the mean and the covariance matrix. They showed that even a small change in the mean can produce a large change in the portfolio position. Thus, the associated risk grows due to the uncertainty of the underlying probability distribution. The relevant keywords in this context are robustness and robust portfolio selection.

Chen, He and Zhang [6] pointed that the assumptions on the distribution are arguably always subjective. Therefore, estimation on the moments of asset returns using the historical data may be considered more objective measurement. Using the knowledge of the mean and the covariance, they introduced (see [6]) analytical solution of the robust portfolio selection based on LPM, VaR or CVaR, as a risk measure.

Lobo and Boyd [11], Costa and Paiva [5], Goldfarb and Iyengar [9], and Lu [12] studied the robust portfolio in the mean variance framework. Instead of the precise information on the mean and the covariance matrix of asset returns, they introduced some types of uncertainties, such as polytopic, box and ellipsoidal uncertainty.

This thesis is outlined as follows: In the first chapter, we introduce risk measures that are often applied to robust portfolio management. In the second chapter, we formulate the corresponding minimization problems for the proposed measures of risk and make further investigation on some special cases of underlying probability distribution. We are particularly interested in the problem of minimizing the worst-case CVaR, or WCVaR, associated with mixture distribution uncertainty, box uncertainty, and ellipsoidal uncertainty in the distributions. In the third chapter, we present the application of WCVaR, introduced by Zhu and Fukushima [17], to robust portfolio optimization. In the last chapter, we discuss the results of numerical applications on portfolio selection performed via the methods proposed in this thesis. Finally, we conclude the results and outline the future directions.

1. Risk Measures

Measures of risk have an important role in optimization under uncertainty, especially in finance and insurance industry. In this chapter, we introduce the most popular risk measures used in risk management, and discuss their fundamental properties.

In the return-risk trade-off analysis, the risk is explicitly quantified by a risk measure that maps the loss to a real number. In general, loss can be expressed as a function $Z = f(\mathbf{x}, \mathbf{y})$ of a decision vector $\mathbf{x} \in \mathscr{X} \subseteq \mathbb{R}^n$ representing portfolio, where $\mathscr{X}$ expresses decision constraints, and a random vector $\mathbf{y} \in \mathscr{Y} \subseteq \mathbb{R}^m$ representing the future values, e.g., interest rates, random rates of return. When $\mathbf{y}$ has a known probability distribution, a random variable Z has its distribution dependent on the choice of $\mathbf{x}$. Therefore, if we want to choose $\mathbf{x}$ within terms of any optimization problem, then we should take into account not just expectations, but also the "riskness" of $\mathbf{x}$.

There is a number of ways that risk can be defined. The important question is, how a suitable risk measure looks like, and what a risk measure might or not might have. In 1999, Artzner et al. [1] presented their essential work on coherent risk measure. They presented the following set of consistency rules for a risk measure ρ mapping a random loss Z to a real number:

(i) *Subadditivity:* For all random losses Z and Y, $\rho(Z+Y) \leq \rho(Z) + \rho(Y)$;

(ii) *Positive homogenity:* For positive constant λ, $\rho(\lambda Z) = \lambda \rho(X)$;

(iii) *Monotonicity:* If $Z \leq Y$ a.s. for each outcome, then $\rho(Z) \leq \rho(Y)$;

(iv) *Translation invariance:* For any constant c, $\rho(Z+c) = \rho(X) + c$.

A risk measure that satisfies the above axioms is called a *coherent risk measure.*

Let random loss Z be defined on some probability space $(\Omega, \mathscr{F}, P)$. In the situation that a probability measure P is ambiguous and characterized as a certain set $\mathscr{P}$, then we generally define the worst-case risk measure ρ_{W} related to ρ as follows:

$$\rho_{\mathrm{W}}(Z) = \sup_{P \in \mathscr{P}} \rho(Z). \tag{1.1}$$

Proposition 1 ([17])**.** *If ρ associated with crisp probability measure P is a coherent risk measure, then the corresponding ρ_W associated with ambiguous probability measure $\mathscr{P}$ remains a coherent risk measure.*

Coherent risk measure, in the sense introduced by Artzner [1], is supposed to be a "good" measure of risk because it has four desirable properties. In this thesis, different risk measures are discussed and they are not necessarily coherent. We are particularly interested in CVaR, that is a well known coherent risk measure, and the worst-case CVaR. By Proposition 1, the worst-case CVaR is also a coherent risk measure.

Throughout the thesis, we also present the results on a class of downside risk measures known as the lower partial moments, introduced by Bawa [2] and Fishburn [8].

1.1 Lower Partial Moments

The lower partial moments are defined as follows

$$\mathrm{LPM}_m(r) = \mathsf{E}\left[(r - X)_+^m\right], \tag{1.2}$$

where $(t)_+ = \max\{t, 0\}$, X is the asset return ($X = -Z$), r is the return on a benchmark index, and m is a parameter, that can take any nonnegative value. Specifically,

- if $m = 0$, then LPM_0 is clearly the probability of the asset return falling below the benchmark index;
- if $m = 1$, then LPM_1 is the expected shortfall of the investment, falling below the benchmark index;
- if $m = 2$, then LPM_2 is almost an analog of the semivariance, but the deviation references to the benchmark instead of the mean.

As we can see, LPM is specified by r and m. The return r is often set to the risk-free rate, or simply to zero. By choosing the degree m an investor can specify the measure of his/her risk attitude. Intuitively, large values of m will penalize large deviations more than low values.

We denote

$$P = \{p | \mathsf{E}_p[\xi] = \boldsymbol{\mu}, \mathrm{Cov}_p[\xi] = \Gamma \succ 0\},$$

where P is the set of probability distributions with mean $\boldsymbol{\mu} \in \mathbb{R}^n$ and covariance matrix $\Gamma \in S^n_{++}$, which is positive semidefinite. We denote $R \sim (\boldsymbol{\mu}, \Gamma)$ to represent the fact that the random vector R belongs to the set whose elements have mean $\boldsymbol{\mu}$ and covariance matrix Γ.

1.1.1 Upper Bounds for the Univariate Cases

In this section, we discuss the moment upper bounds, using the information about the mean and the covariance of the underlying distribution. These bounds lead to robust portfolio optimization models, as we will see later.

As we remarked, $\mathrm{LPM}_0(r)$ measures the probability that a random return falls below the target r. Its upper bound, which is presented in the following lemma, is exactly set by Chebyshev-Cantelli inequality [4].

Lemma 2 ([6]). *It holds that*

$$\begin{aligned} \sup_{X\sim(\mu,\sigma^2)} LPM_0(r) &= \sup_{X\sim(\mu,\sigma^2)} \mathsf{P}\{X \le r\} \\ &= \begin{cases} \frac{1}{1+(r-\mu)^2/\sigma^2}, & \text{if } r < \mu, \\ 1, & \text{if } r \ge \mu. \end{cases} \end{aligned}$$

$\mathrm{LPM}_1(r)$ is the expected shortfall of X below the benchmark r. As we will see later, this measure of risk is highly related to CVaR. In this case, Jensen's inequality is used to derive its upper bound.

Lemma 3 ([6]). *It holds that*

$$\sup_{X\sim(\mu,\sigma^2)} LPM_1(r) = \sup_{X\sim(\mu,\sigma^2)} \mathsf{E}\,[(r-X)_+] = \frac{r-\mu+\sqrt{\sigma^2+(r-\mu)^2}}{2}.$$

A tight upper bound on $\mathrm{LPM}_2(r)$ can be also established by Jensen's inequality.

Lemma 4 ([6]). *It holds that*

$$\sup_{X\sim(\mu,\sigma^2)} LPM_2(r) = \sup_{X\sim(\mu,\sigma^2)} \mathsf{E}\,[(r-X)_+^2] = [(r-\mu)_+]^2+\sigma^2.$$

Proof. Firstly, it holds

$$\sigma^2 = \operatorname{var} X = \mathsf{E}\; X^2 - (\mathsf{E}\; X)^2 = \mathsf{E}\; X^2 - \mu^2,$$

and thus

$$\mathsf{E}\; X^2 = \mu^2+\sigma^2.$$

Further, for any $X\sim(\mu,\sigma^2)$, by Jensen's inequality we have

$$\begin{aligned}
\mathsf{E}\,[(r-X)_+^2] &= \mathsf{E}\,[(r-X)^2] - \mathsf{E}\,[(r-X)_-^2] \\
&\le \mathsf{E}\,[(r-X)^2] - \mathsf{E}\,[(r-X)_-]^2 \\
&= \mathsf{E}\,[r^2-2rX-X^2] - ([\mathsf{E}\; r - \mathsf{E}\; X]_-)^2 \\
&= r^2-2rX-\mathsf{E}\; X^2 - [(r-\mu)_-]^2 \\
&= r^2-2rX+(\mu^2+\sigma^2) - [(r-\mu)_-]^2 \\
&= (r^2-2rX+\mu^2)+\sigma^2 - [(r-\mu)_-]^2 \\
&= (r-\mu)^2+\sigma^2-[(r-\mu)_-]^2 \\
&= \sigma^2-[(r-\mu)_+]^2.
\end{aligned}$$

To show the tightness of the bound, consider a sequence of distributions

$$X_n = \begin{cases} \mu+\frac{\sigma}{\sqrt{n-1}}, & \text{with probability } \frac{n-1}{n}, \\ \mu-\sigma\sqrt{n-1}, & \text{with probability } 1/n. \end{cases}$$

We determine mean and variance of X_n

$$\begin{aligned}
\mathsf{E}\; X_n &= \frac{n-1}{n}\left(\mu+\frac{\sigma}{\sqrt{n-1}}\right)+\frac{1}{n}\left(\mu+\sigma\sqrt{n-1}\right) \\
&= \frac{1}{n}\left(n\mu-\mu+\sigma\sqrt{n-1}\right)+\frac{1}{n}\left(\mu-\sigma\sqrt{n-1}\right) \\
&= \frac{1}{n}\left(n\mu-\mu+\sigma\sqrt{n-1}+\mu-\sigma\sqrt{n-1}\right) \\
&= \frac{1}{n}n\mu \\
&= \mu
\end{aligned}$$

$$
\begin{aligned}
\operatorname{var} X_n &= \mathsf{E}\, X_n^2 - \mathsf{E}\,(X_n)^2 \\
&= \frac{n-1}{n}\left(\mu + \frac{\sigma}{\sqrt{n-1}}\right)^2 + \frac{1}{n}\left(\mu + \sigma\sqrt{n-1}\right)^2 - \mu^2 \\
&= \frac{n-1}{n}\left(\mu^2 + \frac{2\mu\sigma^2}{\sqrt{n-1}} + \frac{\sigma}{n-1}\right) + \frac{1}{n}\left(\mu^2 - 2\mu\sigma\sqrt{n-1} + (n-1)\sigma^2\right) \\
&\quad -\mu^2 \\
&= \frac{1}{n}\left((n-1)\mu^2 + 2\mu\sigma\sqrt{n-1} + \sigma^2\right) + \frac{1}{n}\left(\mu^2 - 2\mu\sigma\sqrt{n-1} + \sigma^2 n - \sigma^2\right) \\
&\quad -\mu^2 \\
&= \frac{1}{n}\left((n\mu^2 + \mu^2 + 2\mu\sigma\sqrt{n-1} + \sigma^2 + \mu^2 - 2\mu\sigma\sqrt{n-1} + \sigma^2 n - \sigma^2\right) - \mu^2 \\
&= \frac{1}{n}\left(n\mu^2 + n\sigma^2\right) - \mu^2 \\
&= \sigma^2 \qquad (1.3)
\end{aligned}
$$

It means that $X_n \sim (\mu, \sigma^2)$, and

$$\mathsf{E}\,[(r - X_n)_+^2] \longrightarrow \sigma^2 + [(r-\mu)_+]^2, \text{as } n \longrightarrow \infty.$$

This indicates that the upper bound is indeed tight, which completes the proof. □

Remark. If $m > 2$, then $\sup_{X \sim (\mu,\sigma^2)} \mathsf{E}\,[(r - X)_+^m] = +\infty$, see [6].

1.2 Value-at-risk

Value-at-risk is one of the most popular risk measures. However, it is unstable and numerical application is difficult when losses does not follow normal distribution. Beyond the treshold amount indicated by this measure, there is no handle on the extent of the losses that might occure. It is incapable of distinguishing between situations when losses are only a little bit worse and those which are enormous. Moreover, VaR is not a coherent risk measure in the sense of Artzner [1]. Despite of all shortcomings, VaR is frequently used in risk management.

In everything that follows, we suppose that random vector $\mathbf{y}$ is governed by a probability measure P on $\mathscr{Y}$ (a Borel measure) that is independent of $\mathbf{x}$. For each $\mathbf{x}$, we denote by $\Psi(\mathbf{x}, \cdot)$ on $\mathbb{R}$ the resulting distribution function for the loss $f(\mathbf{x}, \mathbf{y})$, i.e.,

$$\Psi(\mathbf{x}, \zeta) = \mathsf{P}\{\mathbf{y} | f(\mathbf{x}, \mathbf{y}) \leq \zeta\}. \qquad (1.4)$$

The previous function (1.4) actually represents the probability that $f(\mathbf{x}, \mathbf{y})$ does not exceed a treshold ζ.

Given a confidence level $\alpha \in (0, 1)$ (usually greater than 0.9), the value-at-risk is defined as follows.

Definition 1. *The VaR_α of the loss associated with a decision $\boldsymbol{x}$ is the value*

$$\zeta_\alpha(\boldsymbol{x}) = min\{\zeta \mid \psi(\boldsymbol{x}, \zeta) \geq \alpha\}. \qquad (1.5)$$

Definition 2. *The* VaR_α^+ *("upper"* VaR_α*) of the loss associated with a decision* $\boldsymbol{x}$ *is the value*

$$\zeta_\alpha^+(\boldsymbol{x}) = min\{\zeta \mid \psi(\boldsymbol{x}, \zeta) > \alpha\}. \tag{1.6}$$

Remark. The minimum in (1.5) is achieved because $\psi(\mathbf{x}, \zeta)$ is nondecreasing and right-continuous in ζ. In the situation when $\psi(\mathbf{x}, \zeta)$ is strictly increasing and continuous, $\zeta_\alpha(\mathbf{x})$ is equal to the unique ζ satisfying $\psi(\mathbf{x}, \zeta) = \alpha$. In all other cases, this equation can have no solution or a whole range of solutions.

1.3 Conditional Value-at-risk

A coherent risk measure that quantifies the losses that might occure in the tail is conditional value-at-risk. This risk assessment technique is derived by taking a weighted average between the value-at-risk and losses exceeding the value-at-risk. Therefore, the term is also known as "Mean Excess Loss" or "Tail VaR".

We make a technical assumption that $f(\mathbf{x}, \mathbf{y})$ is continuous in $\mathbf{x}$ and measurable in $\mathbf{y}$. We also assume that $\mathsf{E}\left(|f(\mathbf{x}, \mathbf{y})|\right) < +\infty$ for each $\mathbf{x} \in \mathscr{X}$.

Definition 3. *The* $CVaR_\alpha$ *of the loss associated with a decision* $\boldsymbol{x}$ *is the value*

$$\phi_\alpha(\boldsymbol{x}) = mean\ of\ the\ \alpha - tail\ distribution\ of\ f(\boldsymbol{x}, \boldsymbol{y}),$$

where the distribution in question is the one with distribution function $\psi_\alpha(\boldsymbol{x}, \cdot)$ *defined by*

$$\psi_\alpha(\boldsymbol{x}, \zeta) = \begin{cases} 0, & for\ \zeta < \zeta_\alpha(\boldsymbol{x}) \\ [\psi(\boldsymbol{x}, \zeta) - \alpha]/[1 - \alpha], & for\ \zeta \geq \zeta_\alpha(\boldsymbol{x}). \end{cases} \tag{1.7}$$

Definition 4. *The* $CVaR_\alpha^+$ *("upper"* $CVaR_\alpha$*) of the loss associated with a decision* $\boldsymbol{x}$ *is the value*

$$\psi_\alpha^+ = \mathsf{E}\{f(\boldsymbol{x}, \boldsymbol{y}) | f(\boldsymbol{x}, \boldsymbol{y}) > \zeta_\alpha(\boldsymbol{x})\}, \tag{1.8}$$

whereas the $CVaR_\alpha^-$ *("lower"* $CVaR_\alpha$*) of the loss associated with a decision* $\boldsymbol{x}$ *is the value*

$$\psi_\alpha^- = \mathsf{E}\{f(\boldsymbol{x}, \boldsymbol{y}) | f(\boldsymbol{x}, \boldsymbol{y}) \geq \zeta_\alpha(\boldsymbol{x})\}. \tag{1.9}$$

We note that the conditional expectation in (1.9) is well defined because $\mathsf{P}\{f(\boldsymbol{x}, \boldsymbol{y}) | f(\boldsymbol{x}, \boldsymbol{y}) > \zeta_\alpha(\mathbf{x})\} \geq 1 - \alpha > 0$, but (1.8) only makes sense as long as $\mathsf{P}\{f(\boldsymbol{x}, \boldsymbol{y}) | f(\boldsymbol{x}, \boldsymbol{y} > \zeta_\alpha(\mathbf{x})\} > 0$, which is not assured merely through the assumption that $\alpha \in (0, 1)$. For more details see [16].

1.4 Worst-case VaR and Worst-case CVaR

Instead of assuming the precise knowledge of the distribution of the random vector $\mathbf{y}$, we assume that the density function is only known to belong to a certain set $\mathscr{P}$ of distributions, i.e., $p(\cdot) \in \mathscr{P}$.

According to general definition of the worst-case risk measure (1.1), we define the worst-case VaR and the worst-case CVaR as follows.

Definition 5. *The* $WVaR_\alpha$ *of the loss associated with a decision* $\boldsymbol{x}$ *is the value*

$$WVaR_\alpha(\boldsymbol{x}) = \sup_{p(\cdot) \in \mathscr{P}} VaR_\alpha(\boldsymbol{x}).$$

Definition 6. *The* $WCVaR_\alpha$ *of the loss associated with a decision* $\boldsymbol{x}$ *is the value*

$$WCVaR_\alpha(\boldsymbol{x}) = \sup_{p(\cdot)\in\mathscr{P}} CVaR_\alpha(\boldsymbol{x}).$$

2. Minimization of Risk Measures

In this chapter, we discuss the minimization problem of CVaR and WCVaR. For WCVaR we make further investigation on some special cases of $\mathscr{P}$, formulate the corresponding minimization problems, that can be efficiently solved, and discuss their computational aspects. We are particularly interested in mixture distribution, box uncertainty in discrete distribution and ellipsoidal uncertainty in discrete distribution, which are the most often used uncertainty structures in robust optimization.

2.1 Minimization of CVaR

Suppose that $\boldsymbol{y}$ follows a continuous distribution. As Rockafellar and Uryasev demonstrate [15], the calculation of CVaR can be achieved by minimizing of the following auxiliary function with respect to the variable $\zeta \in \mathbb{R}$:

$$F_\alpha(\boldsymbol{x}, \zeta) = \zeta + \frac{1}{1-\alpha} \int_{\boldsymbol{y} \in \mathbb{R}^m} [f(\boldsymbol{x}, \boldsymbol{y}) - \zeta]^+ p(\boldsymbol{y}) dy, \tag{2.1}$$

where $[t]^+ = \max\{t, 0\}$, and $p(\cdot)$ denotes a density function of $\boldsymbol{y}$. Thus, we have the fundamental minimimization formula specified in the following theorem. For more details and the proof of this theorem see [15].

Theorem 5 (Fundamental minimization formula [15])**.** *As a function of $\zeta \in \mathbb{R}$, $F_\alpha(\boldsymbol{x}, \zeta)$ is finite and convex (hence continuous), with*

$$\phi_\alpha(\boldsymbol{x}) = \min_{\zeta \in \mathbb{R}} F_\alpha(\boldsymbol{x}, \zeta)$$

and moreover

$$\zeta_\alpha(\boldsymbol{x}) = \text{ lower endpoint of } \arg\min_{\zeta \in \mathbb{R}} F_\alpha(\boldsymbol{x}, \zeta),$$

$$\zeta_\alpha^+(\boldsymbol{x}) = \text{ upper endpoint of } \arg\min_{\zeta \in \mathbb{R}} F_\alpha(\boldsymbol{x}, \zeta),$$

where the argmin refers to the set of ζ for which the minimum is attained and in this case has to be a nonempty, closed, bounded interval (perhaps reducing to a single point). In particular, one always has

$$\zeta_\alpha(\boldsymbol{x}) \in \arg\min_{\zeta \in \mathbb{R}} F_\alpha(\boldsymbol{x}, \zeta), \qquad \phi_\alpha(\boldsymbol{x}) = F_\alpha(\boldsymbol{x}, \zeta_\alpha(\boldsymbol{x})).$$

As noticed in [15], Theorem 5 shows the difference between CVaR and VaR, and present the fundamental reason why CVaR is much better behaved than VaR when dependence on a choice of $\boldsymbol{x}$ must be handled. The reason is the fact, that the optimal value in a problem of minimization, in this case $\phi_\alpha(\boldsymbol{x})$, is more agreeable as a function of parameters than is the optimal solution set, which is here the argmin interval with $\zeta_\alpha(\boldsymbol{x})$ as its lower endpoint.

In (2.1) we assume that $\mathbf{y}$ follows a continuous distribution. However, to consider a discrete distribution make sense even for a continuous distribution in CVaR formulation, because we usually use summation to approximate the integral in (2.1). Let S denote the number of sample points. In the following, we assume that the sample space of the random vector $\mathbf{y}$ is given by $\{\mathbf{y}_{(1)}, \mathbf{y}_{(2)}, \ldots, \mathbf{y}_{(S)}\}$, where $\mathsf{P}\{\mathbf{y}_{(k)}\} = \pi_k$ and $\sum_{k=1}^{S} \pi_k = 1$, $\pi_k \geq 0$, $k = 1, \ldots, S$. Further, denote $\boldsymbol{\pi} = (\pi_1, \pi_2, \ldots, \pi_S)^T$ and define

$$G_\alpha(\boldsymbol{x}, \zeta, \boldsymbol{\pi}) = \zeta + \frac{1}{(1-\alpha)} \sum_{k=1}^{S} \pi_k [f(\boldsymbol{x}, \boldsymbol{y}_{(k)}) - \zeta]^+. \tag{2.2}$$

For given $\mathbf{x}$ and $\boldsymbol{\pi}$, Rockafellar and Uryasev [16] defined the corresponding CVaR as

$$\mathrm{CVaR}_\alpha(\mathbf{x}, \boldsymbol{\pi}) = \min_{\zeta \in \mathbb{R}} G_\alpha(\boldsymbol{x}, \zeta, \boldsymbol{\pi}). \tag{2.3}$$

We can formulate minimizing of CVaR as the following minimization program with decision variables $(\mathbf{x}, \mathbf{u}, \zeta) \in \mathbb{R}^n \times \mathbb{R}^S \times \mathbb{R}$:

$$\begin{aligned}
\min \quad & \zeta + \frac{1}{1-\alpha} (\boldsymbol{\pi})^T \boldsymbol{u} && (2.4)\\
s.t. \quad & \boldsymbol{x} \in \mathscr{X}, && (2.5)\\
& u_k \geq f(\boldsymbol{x}, \boldsymbol{y}_{(k)}) - \zeta, \quad k = 1, \ldots, S, && (2.6)\\
& u_k \geq 0, \quad k = 1, \ldots, S, && (2.7)
\end{aligned}$$

where $\mathbf{u} = (u_1, \ldots, u_S)$ is an auxiliary vector utilized to deal with the computation of $[\cdot]_+$ in the original objective function.

If the function $f(\mathbf{x}, \mathbf{y})$ is linear with respect to $\mathbf{x}$ and the set $\mathscr{X}$ is a convex polyhedron, then the problem (2.4) - (2.7) can be solved by a linear programming method.

2.2 Minimization of Worst-case CVaR

In this section, we present the results formulated by Zhu and Fukushima [17]. First of all, we quote the following lemma (minimax theorem), which will be used to formulate the minimization problems in a tractable way.

Lemma 6 ([17])**.** *Suppose that $\mathscr{X}$ and $\mathscr{Y}$ are nonempty compact convex sets in $\mathbb{R}^n$ and $\mathbb{R}^m$, respectively, and the function $\phi(\boldsymbol{x}, \boldsymbol{y})$ is convex in $\boldsymbol{x}$ for any given $\boldsymbol{y}$, and concave in $\boldsymbol{y}$ for any given $\boldsymbol{x}$. Then, we have*

$$\min_{\boldsymbol{x} \in \mathscr{X}} \max_{\boldsymbol{y} \in \mathscr{Y}} \phi(\boldsymbol{x}, \boldsymbol{y}) = \max_{\boldsymbol{y} \in \mathscr{Y}} \min_{\boldsymbol{x} \in \mathscr{X}} \phi(\boldsymbol{x}, \boldsymbol{y})$$

One can find the details and the proof of Lemma 6 for example in [7].

2.2.1 Mixture Distribution

In this section, we assume that about the distribution of $\mathbf{y}$ we only know it belongs to a set of distributions that consists of all mixtures of some predetermined likelihood distributions, i.e.,

$$p(\cdot) \in \mathscr{P}_M = \left\{ \sum_{i=1}^{l} \lambda_i p^i(\cdot) : \sum_{i=1}^{l} \lambda_i = 1, \lambda_i \geq 0, i = 1, \dots, l \right\}, \tag{2.8}$$

where $p^i(\cdot)$ denotes the ith likelihood distribution, and l denotes the number of the likelihood distributions. Denote

$$\Lambda = \left\{ \boldsymbol{\lambda} = (\lambda_1, \dots, \lambda_l) : \sum_{i=1}^{l} \lambda_i = 1, \lambda_i \geq 0, i = 1, \dots, l \right\}. \tag{2.9}$$

Define

$$F_\alpha^i(\boldsymbol{x}, \zeta) = \zeta + \frac{1}{1-\alpha} \int_{\boldsymbol{y} \in \mathbb{R}^m} [f(\boldsymbol{x}, \boldsymbol{y}) - \zeta]^+ p^i(\boldsymbol{y}) d\boldsymbol{y}, i = 1, \dots, l. \tag{2.10}$$

Using Lemma 6, we get the following theorem.

Theorem 7 ([17]). *For each $\boldsymbol{x}$, $WCVaR_\alpha(\boldsymbol{x})$ with respect to $\mathscr{P}_M$ can be computed as*

$$WCVaR_\alpha(\boldsymbol{x}) = \min_{\zeta \in \mathbb{R}} \max_{i \in \mathscr{L}} F_\alpha^i(\boldsymbol{x}, \zeta), \tag{2.11}$$

where $\mathscr{L} = \{1, 2, \dots, l\}$.

Proof. For given $\mathbf{x} \in \mathscr{X}$, we define the following function

$$\begin{aligned} H_\alpha(\boldsymbol{x}, \zeta, \boldsymbol{\lambda}) &= \zeta + \frac{1}{1-\alpha} \int_{\boldsymbol{y} \in \mathbb{R}^m} [f(\boldsymbol{x}, \boldsymbol{y}) - \zeta]^+ \left[\sum_{i=1}^{l} \lambda_i p^i(\boldsymbol{y}) \right] d\boldsymbol{y} \\ &= \sum_{i=1}^{l} \lambda_i F_\alpha^i(\boldsymbol{x}, \zeta) \end{aligned} \tag{2.12}$$

where $\boldsymbol{\lambda} \in \Lambda$, and the set Λ is specified as (2.9). The function $H_\alpha(\boldsymbol{x}, \zeta, \boldsymbol{\lambda})$ is convex in ζ and concave in $\boldsymbol{\lambda}$ [16]. Moreover, $\min_{\zeta \in \mathbb{R}} H_\alpha(\boldsymbol{x}, \zeta, \boldsymbol{\lambda})$ is a continuous function with respect to $\boldsymbol{\lambda}$. By the definition of $\mathrm{WCVaR}_\alpha(\boldsymbol{x})$, and the fact that Λ is a compact set, we can write

$$\begin{aligned} \mathrm{WCVaR}_\alpha(\boldsymbol{x}) &= \max_{\boldsymbol{\lambda} \in \Lambda} \min_{\zeta \in \mathbb{R}} H_\alpha(\boldsymbol{x}, \zeta, \boldsymbol{\lambda}) \\ &= \max_{\boldsymbol{\lambda} \in \Lambda} \min_{\zeta \in \mathbb{R}} \sum_{i=1}^{l} \lambda_i F_\alpha^i(\boldsymbol{x}, \zeta) \end{aligned} \tag{2.13}$$

As Rockafellar and Uryasev proved [16], for fixed $\mathbf{x}$ and each i, the optimal solution set of $\min_{\zeta \in \mathbb{R}} F_\alpha^i(\boldsymbol{x}, \zeta)$ is a nonempty, closed, and bounded interval. Thus, we denote

$$[\underline{\alpha}_i^*, \bar{\alpha}_i^*] = \arg\min_{\zeta\in\mathbb{R}} F_\alpha^i(\boldsymbol{x},\zeta), i = 1,\ldots,l.$$

Suppose that $g_1(t)$ and $g_2(t)$ are two convex functions defined on $\mathbb{R}$. Let the nonempty, closed, and bounded intervals $[\underline{t}_1^*, \bar{t}_1^*]$ and $[\underline{t}_2^*, \bar{t}_2^*]$ denote their sets of minima. It holds that for any $\beta_1 \geq 0$ and $\beta_2 \geq 0$ such that $\beta_1 + \beta_2 > 0$, $\beta_1 g_1(t) + \beta_2 g_2(t)$ is also a convex function, and its set of minima must lie in the nonempty, closed, and bounded interval $[\min\{\underline{t}_1^*, \underline{t}_2^*\}, \max\{\bar{t}_1^*, \bar{t}_2^*\}]$. By this fact, we get

$$\arg\min_{\zeta\in\mathbb{R}} H_\alpha(\boldsymbol{x},\zeta,\boldsymbol{\lambda}) \subseteq \mathscr{A}, \forall\boldsymbol{\lambda}\in\Lambda,$$

where $\mathscr{A}$ is the nonempty, closed, and bounded interval given by

$$\mathscr{A} = [\min_{i\in\mathscr{L}} \underline{\alpha}_i^*, \max_{i\in\mathscr{L}} \bar{\alpha}_i^*],$$

which implies

$$\min_{\zeta\in\mathbb{R}} H_\alpha(\boldsymbol{x},\zeta,\boldsymbol{\lambda}) = \min_{\zeta\in\mathscr{A}} H_\alpha(\boldsymbol{x},\zeta,\boldsymbol{\lambda}).$$

Thus, by Lemma 6

$$\begin{aligned}\max_{\boldsymbol{\lambda}\in\Lambda}\min_{\zeta\in\mathbb{R}} H_\alpha(\boldsymbol{x},\zeta,\boldsymbol{\lambda}) &= \max_{\boldsymbol{\lambda}\in\Lambda}\min_{\zeta\in\mathscr{A}} H_\alpha(\boldsymbol{x},\zeta,\boldsymbol{\lambda}) \\ &= \min_{\zeta\in\mathscr{A}}\max_{\boldsymbol{\lambda}\in\Lambda} H_\alpha(\boldsymbol{x},\zeta,\boldsymbol{\lambda}).\end{aligned} \tag{2.14}$$

Obviously, it holds

$$\min_{\zeta\in\mathscr{A}}\max_{\boldsymbol{\lambda}\in\Lambda} H_\alpha(\boldsymbol{x},\zeta,\boldsymbol{\lambda}) \geq \inf_{\zeta\in\mathbb{R}}\max_{\boldsymbol{\lambda}\in\Lambda} H_\alpha(\boldsymbol{x},\zeta,\boldsymbol{\lambda}) \tag{2.15}$$

By (2.14), (2.15), and the well-known result on the min-max inequality

$$\inf_{\zeta\in\mathbb{R}}\max_{\boldsymbol{\lambda}\in\Lambda} H_\alpha(\boldsymbol{x},\zeta,\boldsymbol{\lambda}) \geq \max_{\boldsymbol{\lambda}\in\Lambda}\min_{\zeta\in\mathbb{R}} H_\alpha(\boldsymbol{x},\zeta,\boldsymbol{\lambda}),$$

we get

$$\max_{\boldsymbol{\lambda}\in\Lambda}\min_{\zeta\in\mathbb{R}} H_\alpha(\boldsymbol{x},\zeta,\boldsymbol{\lambda}) = \min_{\zeta\in\mathbb{R}}\max_{\boldsymbol{\lambda}\in\Lambda} H_\alpha(\boldsymbol{x},\zeta,\boldsymbol{\lambda}).$$

Therefore, we can write

$$\begin{aligned}\mathrm{WCVaR}_\alpha(\boldsymbol{x}) &= \min_{\zeta\in\mathbb{R}}\max_{\boldsymbol{\lambda}\in\Lambda} H_\alpha(\boldsymbol{x},\zeta,\boldsymbol{\lambda}) \\ &= \min_{\zeta\in\mathbb{R}}\max_{\boldsymbol{\lambda}\in\Lambda} \sum_{i=1}^{l} \lambda_i F_\alpha^i(\boldsymbol{x},\zeta).\end{aligned} \tag{2.16}$$

Now we only need to verify that the righ-hand sides of (2.11) and (2.16) are equivalent. The right-hand side of (2.16) can be written as the following optimization problem

$$\min_{(\zeta,\theta)\in\mathbb{R}\times\mathbb{R}} \left\{\theta : \sum_{i=1}^{l} \lambda_i F_\alpha^i(\boldsymbol{x},\zeta) \leq \theta, \forall\boldsymbol{\lambda}\in\Lambda\right\}. \tag{2.17}$$

By the specification of the set Λ (2.9), it is clear that any feasible solution of (2.17) satisfies

$$F^i_\alpha(\boldsymbol{x}, \zeta) \le \theta, i = 1, \ldots, l. \tag{2.18}$$

On the other hand, if (2.18) holds, then for any $\boldsymbol{\lambda} \in \Lambda$, we have

$$\sum_{i=1}^{l} \lambda_i F^i_\alpha(\boldsymbol{x}, \zeta) \le \sum_{i=1}^{l} \lambda_i \theta = \theta.$$

Thus, we can see that the problem (2.17) is equivalent to

$$\min_{(\zeta,\theta)\in\mathbb{R}\times\mathbb{R}} \left\{\theta : F^i_\alpha(\boldsymbol{x}, \zeta) \le \theta, i = 1, \ldots, l\right\},$$

which is in fact the righ-hand side of (2.11) written as an optimization problem. This completes the proof. □

Denote

$$F^{\mathscr{L}}_\alpha(\boldsymbol{x}, \zeta) = \max_{i\in\mathscr{L}} F^i_\alpha(\boldsymbol{x}, \zeta). \tag{2.19}$$

By Theorem 7, we get the following corollary.

Corollary 8 ([17])**.** *Minimizing $WCVaR_\alpha(\boldsymbol{x})$ over $\mathscr{X}$ can be achieved by minimizing $F^{\mathscr{L}}_\alpha(\boldsymbol{x}, \zeta)$ over $\mathscr{X} \times \mathbb{R}$, i.e.,*

$$\min_{\boldsymbol{x}\in\mathscr{X}} WCVaR_\alpha(\boldsymbol{x}) = \min_{(\boldsymbol{x},\zeta)\in\mathscr{X}\times\mathbb{R}} F^{\mathscr{L}}_\alpha(\boldsymbol{x}, \zeta). \tag{2.20}$$

More specifically, if $(\boldsymbol{x}^*, \zeta^*)$ attains the right-hand side minimum, then $\boldsymbol{x}^*$ attains the left-hand side minimum, and ζ^* attains the minimum of $F^{\mathscr{L}}_\alpha(\boldsymbol{x}^*, \zeta)$, and vice versa.

As Rockafellar and Uryasev demonstrate [16], the function $F_\alpha(\boldsymbol{x}, \zeta)$ defined by 2.1 is convex in $(\mathbf{x}, \zeta)$ if the function $f(\mathbf{x}, \mathbf{y})$ is convex in $\mathbf{x}$. It holds that the function $g(t) = \max\{g_1(t), g_2(t)\}$ is convex whenever both $g_1(t)$ and $g_2(t)$ are convex. Thus, we can see that if $f(\mathbf{x}, \mathbf{y})$ is convex in $\mathbf{x}$, then $F^{\mathscr{L}}_\alpha(\boldsymbol{x}, \zeta)$ is convex in $(\mathbf{x}, \zeta)$. Moreover, if $f(\mathbf{x}, \mathbf{y})$ is convex in $\mathbf{x}$ and $\mathscr{X}$ is a convex set, then the problem of WCVaR minimization is a convex program.

Using Theorem 7 and Corollary (8), we get that the WCVaR minimization problem is equivalent to

$$\min_{(\boldsymbol{x},\zeta,\theta)\in\mathscr{X}\times\mathbb{R}\times\mathbb{R}} \left\{\theta : \zeta + \frac{1}{1-\alpha}\int_{\boldsymbol{y}\in\mathbb{R}^m} [f(\boldsymbol{x}, \boldsymbol{y}) - \zeta]^+ p^i(\boldsymbol{y}) d\boldsymbol{y} \le \theta, i = 1, \ldots, l\right\}, \tag{2.21}$$

which is more tractable problem in comparison with the original one. The calculation of the integral is probably the most difficult part in this computation. Monte Carlo simulation is one of the most popular and efficient approximation methods used to deal with this complexity. Rockefellar and Uryasev [16] use this method to approximate $\tilde{F}_\alpha(\boldsymbol{x}, \zeta)$ as follows

$$\tilde{F}_\alpha(\boldsymbol{x},\zeta) = \zeta + \frac{1}{S(1-\alpha)} \sum_{k=1}^{S} [f(\boldsymbol{x}, \boldsymbol{y}_{(k)}) - \zeta]^+, \tag{2.22}$$

where $\mathbf{y}_{(k)}$ denotes the kth sample point that is generated by simple random sampling according to density function $p(\cdot)$ of $\mathbf{y}$, and S denotes the number of sample points. When the number of sample points used in approximation is large enough, then the approximation accuracy (or convergence) is guaranteed by the law of large numbers.

Remark. We note that the function $G_\alpha(\boldsymbol{x},\zeta,\boldsymbol{\pi})$ (2.2) is equal to the function $\tilde{F}_\alpha(\mathbf{x},\zeta)$ (2.22), if π_k is equal to $1/S$.

In the following, we replace the integral in (2.21) with the summation used in (2.22)

$$\min_{(\boldsymbol{x},\zeta,\theta)\in\mathscr{X}\times\mathbb{R}\times\mathbb{R}} \left\{ \theta : \zeta + \frac{1}{S^i(1-\alpha)} \sum_{k=1}^{S^i} [f(\boldsymbol{x}, \boldsymbol{y}_{(k)})^i - \zeta]^+ \leq \theta, i = 1, \ldots, l \right\}, \tag{2.23}$$

where $\mathbf{y}^i_{(k)}$ denotes the kth sample point with respect to the ith likelihood distribution $p^i(\cdot)$, and S^i denotes the number of sample points.

In general, the problem (2.21) can be formulated using the approximation as

$$\min_{(\boldsymbol{x},\zeta,\theta)\in\mathscr{X}\times\mathbb{R}\times\mathbb{R}} \left\{ \theta : \zeta + \frac{1}{(1-\alpha)} \sum_{k=1}^{S^i} \pi^i_k [f(\boldsymbol{x}, \boldsymbol{y}_{(k)})^i - \zeta]^+ \leq \theta, i = 1, \ldots, l \right\}, \tag{2.24}$$

where π^i_k denotes the probability with respect to the ith likelihood distribution $p^i(\cdot)$ according to the kth sample point.

Remark. Obviously, if π^i_k is equal to $1/S^i$ for all k, then (2.24) can be reducing to problem (2.23).

This general form (2.24) of the optimization problem can be reformulated as the following problem with decision variables $(\mathbf{x},\mathbf{u},\zeta,\theta) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R} \times \mathbb{R}$:

$$\begin{aligned}
\min \quad & \theta && (2.25)\\
s.t. \quad & \boldsymbol{x} \in \mathscr{X}, && (2.26)\\
& \zeta + \frac{1}{1-\alpha} (\boldsymbol{\pi}^i)^T \boldsymbol{u}^i \leq \theta, \quad i = 1, \ldots, l, && (2.27)\\
& u^i_k \geq f(\boldsymbol{x}, \boldsymbol{y}^i_{(k)}) - \zeta, \quad k = 1, \ldots, S^i, i = 1, \ldots, l, && (2.28)\\
& u^i_k \geq 0, k = 1, \ldots, S^i, \quad i = 1, \ldots, l. && (2.29)
\end{aligned}$$

where $\boldsymbol{\pi}^i = (\pi^i_1, \ldots, \pi^i_{S^i})$, and $\mathbf{u} = (\mathbf{u}^1; \mathbf{u}^2; \ldots; \mathbf{u}^l) \in \mathbb{R}^m$ is an auxiliary vector, where $m = \sum_{i=1}^{l} S^i$.

If the function $f(\mathbf{x},\mathbf{y})$ is linear with respect to $\mathbf{x}$ and the set $\mathscr{X}$ is a convex polyhedron, then the problem (2.25) - (2.29) is a linear program.

Remark. Especially if $l = 1$, then the problem (2.25) - (2.29) is exactly that of Rockafellar and Uryasev [16] with $\pi^i_k = 1/S^1$.

2.2.2 Discrete Distribution

In this section, we assume that $\mathbf{y}$ follows a discrete distribution. We are particularly interested in minimizing of WCVaR under box uncertainty and ellipsoidal uncertainty. These two types of uncertainty sets are easy to be specified and the optimization problem can be formulated in a tractable way. In the previous chapter, we defined $\mathrm{WCVaR}_\alpha(\mathbf{x})$ for the general distribution case. In the case of discrete distribution, we denote $\mathscr{P}$ as $\mathscr{P}_\pi$, that we may identify as a subset of $\mathbb{R}^S$. By the formula (2.3), WCVaR for fixed $\mathbf{x} \in \mathscr{X}$ with respect to $\mathscr{P}_\pi$ is then defined as

$$\mathrm{WCVaR}_\alpha(\mathbf{x}) = \sup_{\boldsymbol{\pi} \in \mathscr{P}_\pi} \mathrm{CVaR}_\alpha(\mathbf{x}, \boldsymbol{\pi}),$$

and it is equivalent to

$$\mathrm{WCVaR}_\alpha(\mathbf{x}) = \sup_{\boldsymbol{\pi} \in \mathscr{P}_\pi} \min_{\zeta \in \mathbb{R}} G_\alpha(\boldsymbol{x}, \zeta, \boldsymbol{\pi}).$$

Using Lemma 6, we get the following theorem. The proof of this theorem is given for example in [17].

Theorem 9 ([17])**.** *Suppose that $\mathscr{P}_\pi$ is a compact convex set. Then, for each $\boldsymbol{x}$, we have*

$$WCVaR_\alpha(\boldsymbol{x}) = \min_{\zeta \in \mathbb{R}} \max_{\pi \in \mathscr{P}_\pi} G_\alpha(\boldsymbol{x}, \zeta, \boldsymbol{\pi}).$$

By Theorem 9, if $\mathscr{P}_\pi$ is a compact convex set, the minimization problem of $\mathrm{WCVaR}_\alpha(\mathbf{x})$ over $\mathscr{X}$ can be also formulated as the following problem with decision variables $(\mathbf{x}, \mathbf{u}, \zeta, \theta) \in \mathbb{R}^n \times \mathbb{R}^S \times \mathbb{R} \times \mathbb{R}$:

$$\begin{aligned}
\min \quad & \theta && (2.30)\\
s.t. \quad & \boldsymbol{x} \in \mathscr{X}, && (2.31)\\
& \max_{\boldsymbol{\pi} \in \mathscr{P}_\pi} \zeta + \frac{1}{1-\alpha} \boldsymbol{\pi}^T \boldsymbol{u} \le \theta, && (2.32)\\
& u_k \ge f(\mathbf{x}, \mathbf{y}_{(k)}) - \zeta, \quad k = 1, \ldots, S, && (2.33)\\
& u_k \ge 0, \quad k = 1, \ldots, S. && (2.34)
\end{aligned}$$

Problem (2.30) - (2.34) includes the max operation in the constraints and thus it is not suitable for numerical application. If $f(\mathbf{x}, \mathbf{y})$ is linear in $\mathbf{x}$ and $\mathscr{X}$ is a convex polyhedron, then under box uncertainty in distribution, this problem can be formulated as a linear program, and under ellipsoidal uncertainty in distribution, as a second-order cone program [17].

2.2.3 Box Uncertainty in Discrete Distribution

Suppose that $\boldsymbol{\pi}$ belongs to a box, i.e.,

$$\boldsymbol{\pi} \in \mathscr{P}_\pi^B = \{\boldsymbol{\pi} : \boldsymbol{\pi} = \boldsymbol{\pi}^0 + \boldsymbol{\eta}, \mathbf{e}^T \boldsymbol{\eta} = 0, \underline{\boldsymbol{\eta}} \le \boldsymbol{\eta} \le \bar{\boldsymbol{\eta}}\}, \tag{2.35}$$

where $\boldsymbol{\pi}^0$ denotes nominal distribution (the most likely distribution), $\mathbf{e}$ denotes the vector of ones, and $\underline{\boldsymbol{\eta}}$ and $\bar{\boldsymbol{\eta}}$ are given constant vectors. We can see that the constraint $\boldsymbol{e}^T\boldsymbol{\eta} = 0$ ensures $\boldsymbol{\pi}$ to be a probability distribution, and the nonnegativity constraint $\boldsymbol{\pi} \geq 0$ is included in the box constraints $\underline{\boldsymbol{\eta}} \leq \boldsymbol{\eta} \leq \bar{\boldsymbol{\eta}}$.

It holds

$$\zeta + \frac{1}{1-\alpha}\boldsymbol{\pi}^T\boldsymbol{u} = \zeta + \frac{1}{1-\alpha}(\boldsymbol{\pi}^0 + \boldsymbol{\eta})^T\boldsymbol{u} = \zeta + \frac{1}{1-\alpha}(\boldsymbol{\pi}^0)^T\boldsymbol{u} + \frac{1}{1-\alpha}(\boldsymbol{\eta})^T\boldsymbol{u},$$

and thus we get

$$\max_{\boldsymbol{\pi}\in\mathscr{P}_\pi^B} \zeta + \frac{1}{1-\alpha}\boldsymbol{\pi}^T\boldsymbol{u} = \zeta + \frac{1}{1-\alpha}(\boldsymbol{\pi}^0)^T\boldsymbol{u} + \frac{\gamma^*(\boldsymbol{u})}{1-\alpha},$$

where $\gamma^*(\boldsymbol{u})$ denotes the optimal value of the following linear program:

$$\max_{\boldsymbol{\eta}\in\mathbb{R}^S}\{\boldsymbol{u}^T\boldsymbol{\eta} : \boldsymbol{e}^T\boldsymbol{\eta} = 0, \underline{\boldsymbol{\eta}} \leq \boldsymbol{\eta} \leq \bar{\boldsymbol{\eta}}\}. \tag{2.36}$$

The dual program of (2.36) is given by

$$\min_{(z,\boldsymbol{\xi},\boldsymbol{\omega})\in\mathbb{R}\times\mathbb{R}^S\times\mathbb{R}^S}\{\bar{\boldsymbol{\eta}}^T\boldsymbol{\xi} + \underline{\boldsymbol{\eta}}^T\boldsymbol{\omega} : \boldsymbol{e}z + \boldsymbol{\xi} + \boldsymbol{\omega} = \boldsymbol{u}, \boldsymbol{\xi} \geq 0, \boldsymbol{\omega} \leq 0\}. \tag{2.37}$$

By the previous discussion, we get the following minimization problem with decision variables $(\mathbf{x}, \mathbf{u}, z, \boldsymbol{\xi}, \boldsymbol{\omega}, \zeta, \theta) \in \mathbb{R}^n \times \mathbb{R}^S \times \mathbb{R} \times \mathbb{R}^S \times \mathbb{R}^S \times \mathbb{R} \times \mathbb{R}$:

$$
\begin{aligned}
\min\quad & \theta && (2.38)\\
s.t.\quad & \boldsymbol{x} \in \mathscr{X}, && (2.39)\\
& \zeta + \frac{1}{1-\alpha}(\boldsymbol{\pi}^0)^T\boldsymbol{u} + \frac{1}{1-\alpha}(\bar{\boldsymbol{\eta}}^T\boldsymbol{\xi} + \underline{\boldsymbol{\eta}}^T\boldsymbol{\omega}) \leq \theta, && (2.40)\\
& \boldsymbol{e}z + \boldsymbol{\xi} + \boldsymbol{\omega} = \boldsymbol{u}, && (2.41)\\
& \boldsymbol{\xi} \geq 0, \boldsymbol{\omega} \leq 0, && (2.42)\\
& u_k \geq f(\mathbf{x}, \mathbf{y}_{(k)}) - \zeta, \quad k = 1, \ldots, S, && (2.43)\\
& u_k \geq 0, \quad k = 1, \ldots, S. && (2.44)
\end{aligned}
$$

Remark. If $\underline{\boldsymbol{\eta}} = \bar{\boldsymbol{\eta}} = 0$, then the problem (2.38) - (2.44) reduces to the original CVaR minimization problem.

If the function $f(\mathbf{x}, \mathbf{y})$ is convex with respect to $\mathbf{x}$ and $\mathscr{X}$ is a convex set, then (2.38) - (2.44) is a convex program. Moreover, if $f(\mathbf{x}, \mathbf{y})$ is linear in $\mathbf{x}$ and $\mathscr{X}$ is a convex polyhedron, then the problem is a linear program.

Proposition 10 ([17]). *If $(\boldsymbol{x}^*, \boldsymbol{u}^*, z^*, \xi^*, \omega^*, \zeta^*, \theta^*)$ solves (2.38) - (2.44), then $(\boldsymbol{x}^*, \boldsymbol{u}^*, \zeta^*, \theta^*)$ solves (2.30) - (2.34) with $\mathscr{P}_\pi = \mathscr{P}_\pi^B$. Conversely, if $(\tilde{\boldsymbol{x}}^*, \tilde{\boldsymbol{u}}^*, \tilde{\zeta}^*, \tilde{\theta}^*)$ solves (2.30) - (2.34) with $\mathscr{P}_\pi = \mathscr{P}_\pi^B$, then $(\tilde{\boldsymbol{x}}^*, \tilde{\boldsymbol{u}}^*, \tilde{z}^*, \tilde{\xi}^*, \tilde{\omega}^*, \tilde{\zeta}^*, \tilde{\theta}^*)$ solves (2.38) - (2.44), where $(\tilde{z}^*, \tilde{\xi}^*, \tilde{\omega}^*)$ is an optimal solution to (2.38) with $\boldsymbol{u} = \tilde{\boldsymbol{u}}^*$.*

Proof. Let $(\mathbf{x}^*, \mathbf{u}^*, z^*, \xi^*, \omega^*, \zeta^*, \theta^*)$ is an optimal solution to (2.38) - (2.44). It holds that (z^*, ξ^*, ω^*) is feasible to (2.37) with $\mathbf{u} = \mathbf{u}^*$. By the duality theorem of linear programming, we have

$$\gamma^*(\mathbf{u}^*) \leq \bar{\boldsymbol{\eta}}^T \boldsymbol{\xi}^* + \underline{\boldsymbol{\eta}}^T \boldsymbol{\omega}^*.$$

Therefore, we get

$$\begin{aligned} \max_{\boldsymbol{\pi} \in \mathscr{P}_\pi^B} \zeta^* + \frac{1}{1-\alpha} \boldsymbol{\pi}^T \boldsymbol{u}^* &= \zeta^* + \frac{1}{1-\alpha} (\boldsymbol{\pi}^0)^T \boldsymbol{u}^* + \frac{\gamma^*(\boldsymbol{u}^*)}{1-\alpha} \\ &\leq \zeta^* + \frac{1}{1-\alpha} (\boldsymbol{\pi}^0)^T \boldsymbol{u}^* + \frac{1}{1-\alpha} (\bar{\boldsymbol{\eta}}^T \boldsymbol{\xi}^* + \underline{\boldsymbol{\eta}}^T \boldsymbol{\omega}^*) \\ &\leq \theta^*. \end{aligned} \tag{2.45}$$

The last inequality, together with constraints (2.39), (2.43), and (2.44), implies that $(\mathbf{x}^*, \mathbf{u}^*, \zeta^*, \theta^*)$ is feasible to (2.30) - (2.34) with $\mathscr{P}_\pi = \mathscr{P}_\pi^B$.

Now we need to show that $(\mathbf{x}^*, \mathbf{u}^*, \zeta^*, \theta^*)$ is an optimal solution. Suppose that $(\mathbf{x}^*, \mathbf{u}^*, \zeta^*, \theta^*)$ is not an optimal solution to (2.30) - (2.34) with $\mathscr{P}_\pi = \mathscr{P}_\pi^B$. It means, there exist the other optimal solution $(\bar{\mathbf{x}}^*, \bar{\mathbf{u}}^*, \bar{\zeta}^*, \bar{\theta}^*)$ to (2.30) - (2.34) such that $\bar{\theta}^* < \theta^*$.

Let $(\bar{z}^*, \bar{\xi}^*, \bar{\omega}^*)$ be an optimal solution to (2.37) with $\bar{\mathbf{u}} = \bar{\mathbf{u}}^*$. By the duality theorem of linear programming, we get

$$\begin{aligned} \bar{\zeta}^* + \frac{1}{1-\alpha} (\boldsymbol{\pi}^0)^T \bar{\boldsymbol{u}}^* + \frac{1}{1-\alpha} (\bar{\boldsymbol{\eta}}^T \bar{\boldsymbol{\xi}}^* + \underline{\boldsymbol{\eta}}^T \bar{\boldsymbol{\omega}}^*) &= \bar{\zeta}^* + \frac{1}{1-\alpha} (\boldsymbol{\pi}^0)^T \bar{\boldsymbol{u}}^* + \frac{\gamma^*(\bar{\boldsymbol{u}}^*)}{1-\alpha} \\ &= \max_{\boldsymbol{\pi} \in \mathscr{P}_\pi^B} \bar{\zeta}^* + \frac{1}{1-\alpha} \boldsymbol{\pi}^T \bar{\boldsymbol{u}}^* \\ &= \bar{\theta}^*, \end{aligned} \tag{2.46}$$

which, together with the constraints (2.31), (2.33), (2.34), and (2.37), implies that $(\bar{\mathbf{x}}^*, \bar{\mathbf{u}}^*, \bar{z}^*, \bar{\xi}^*, \bar{\omega}^*, \bar{\zeta}^*, \bar{\theta}^*)$ is feasible to (2.38) - (2.44). Due to $\bar{\theta}^* < \theta^*$, this contradicts our assumption that $(\mathbf{x}^*, \mathbf{u}^*, z^*, \xi^*, \omega^*, \zeta^*, \theta^*)$ is an optimal solution to (2.38) - (2.44). Thus $(\mathbf{x}^*, \mathbf{u}^*, \zeta^*, \theta^*)$ is an optimal solution to (2.30) - (2.34) with $\mathscr{P}_\pi = \mathscr{P}_\pi^B$.

On the other hand, let $(\tilde{\mathbf{x}}^*, \tilde{\mathbf{u}}^*, \tilde{\zeta}^*, \tilde{\theta}^*)$ be an optimal solution to (2.30) - (2.34) with $\mathscr{P}_\pi = \mathscr{P}_\pi^B$, and let $(\tilde{z}^*, \tilde{\xi}^*, \tilde{\omega}^*)$ denote an optimal solution to (2.37) with $\mathbf{u} = \tilde{\mathbf{u}}^*$. Then $(\tilde{\mathbf{x}}^*, \tilde{\mathbf{u}}^*, \tilde{z}^*, \tilde{\xi}^*, \tilde{\omega}^*, \tilde{\zeta}^*, \tilde{\theta}^*)$ must solve (2.38) - (2.44). Otherwise, there exist an optimal solution $(\bar{\mathbf{x}}^*, \bar{\mathbf{u}}^*, \bar{z}^*, \bar{\xi}^*, \bar{\omega}^*, \bar{\zeta}^*, \bar{\theta}^*)$ of (2.38) - (2.44) such that $\bar{\theta}^* < \tilde{\theta}^*$. By the first part of the proof, $(\bar{\mathbf{x}}^*, \bar{\mathbf{u}}^*, \bar{\zeta}^*, \bar{\theta}^*)$ is an optimal solution to (2.30) - (2.34), which contradicts the assumption that $(\tilde{\mathbf{x}}^*, \tilde{\mathbf{u}}^*, \tilde{\zeta}^*, \tilde{\theta}^*)$ is an optimal solution to (2.30) - (2.34) because $\bar{\theta}^* < \tilde{\theta}^*$. This completes the proof. □

2.2.4 Ellipsoidal Uncertainty in Discrete Distribution

Suppose that $\boldsymbol{\pi}$ belongs to an ellipsoid, i.e.,

$$\boldsymbol{\pi} \in \mathscr{P}_\pi^E = \{\boldsymbol{\pi} : \boldsymbol{\pi} = \boldsymbol{\pi}^0 + A\boldsymbol{\eta}, \boldsymbol{e}^T A\boldsymbol{\eta} = 0, \boldsymbol{\pi}^0 + A\boldsymbol{\eta} \geq 0, \|\boldsymbol{\eta}\| \leq 1\}, \tag{2.47}$$

where $\|\boldsymbol{\eta}\| = \sqrt{\boldsymbol{\eta}^T\boldsymbol{\eta}}$, $\boldsymbol{\pi}^0$ denotes a nominal distribution that is the center of the ellipsoid, and $A \in \mathbb{R}^{S\times S}$ is the scaling matrix of the ellipsoid. The conditions $\boldsymbol{e}^T A\boldsymbol{\eta} = 0$ and $\boldsymbol{\pi}^0 + A\boldsymbol{\eta} \geq 0$ ensure that $\boldsymbol{\pi}$ is a probability distribution.

By (2.47), it holds

$$\zeta + \frac{1}{1-\alpha}\boldsymbol{\pi}^T\boldsymbol{u} = \zeta + \frac{1}{1-\alpha}(\boldsymbol{\pi}^0 + A\boldsymbol{\eta})^T\boldsymbol{u} = \zeta + \frac{1}{1-\alpha}(\boldsymbol{\pi}^0)^T\boldsymbol{u} + \frac{1}{1-\alpha}A(\boldsymbol{\eta})^T\boldsymbol{u}.$$

Thus, using the constraint (2.32) we get

$$\max_{\boldsymbol{\pi}\in\mathscr{P}_\pi^B} \zeta + \frac{1}{1-\alpha}\boldsymbol{\pi}^T\boldsymbol{u} = \zeta + \frac{1}{1-\alpha}(\boldsymbol{\pi}^0)^T\boldsymbol{u} + \frac{\gamma^*(\boldsymbol{u})}{1-\alpha},$$

where $\gamma^*(\boldsymbol{u})$ denotes the optimal value of the following convex program:

$$\max_{\boldsymbol{\eta}\in\mathbb{R}^S}\{\boldsymbol{u}^T A\boldsymbol{\eta} : \boldsymbol{e}^T A\boldsymbol{\eta} = 0, \boldsymbol{\pi}^0 + A\boldsymbol{\eta} \geq 0, \|\boldsymbol{\eta}\| \leq 1\}, \tag{2.48}$$

The dual of (2.48) is the second-order cone program (see [10])

$$\min_{(\gamma,\boldsymbol{\omega},\boldsymbol{\xi},z)\in\mathbb{R}\times\mathbb{R}^S\times\mathbb{R}^S\times\mathbb{R}}\{\gamma + (\boldsymbol{\pi}^0)^T\omega : -\boldsymbol{\xi} - A^T\boldsymbol{\omega} + A^T\boldsymbol{e}z = A^T\boldsymbol{u}, \|\boldsymbol{\xi}\| \leq \gamma, \boldsymbol{\omega} \geq 0\}. \tag{2.49}$$

Under some mild condition, such as the existence of interior feasible points for both (2.48) and (2.49), the zero duality gap is guaranteed by the strong conic duality theorem [10]. In this case, by using a similar argument to Proposition 10, we can equivalently formulate (2.30) - (2.34), with $\mathscr{P}_\pi = \mathscr{P}_\pi^E$ as the following minimization problem with decision variables $(\mathbf{x}, \mathbf{u}, \boldsymbol{\gamma}, \boldsymbol{\omega}, \boldsymbol{\xi}, z, \zeta, \theta) \in \mathbb{R}^n \times \mathbb{R}^S \times \mathbb{R} \times \mathbb{R}^S \times \mathbb{R}^S \times \mathbb{R} \times \mathbb{R} \times \mathbb{R}$:

$$\begin{aligned}
\min \quad & \theta && (2.50)\\
s.t. \quad & \boldsymbol{x} \in \mathscr{X}, && (2.51)\\
& \zeta + \frac{1}{1-\alpha}(\boldsymbol{\pi}^0)^T\boldsymbol{u} + \frac{1}{1-\alpha}[\gamma + (\boldsymbol{\pi}^0)^T\boldsymbol{\omega}] \leq \theta, && (2.52)\\
& -\boldsymbol{\xi} - A^T\boldsymbol{\omega} + A^T\boldsymbol{e}z = A^T\boldsymbol{u}, && (2.53)\\
& \|\boldsymbol{\xi}\| \leq \gamma, \boldsymbol{\omega} \geq 0, && (2.54)\\
& u_k \geq f(\mathbf{x}, \mathbf{y}_{(k)}) - \zeta, \quad k = 1, \ldots, S, && (2.55)\\
& u_k \geq 0, \quad k = 1, \ldots, S. && (2.56)
\end{aligned}$$

Remark. If $A = \mathbf{0}$, then the problem (2.50) - (2.56) reduces to the original CVaR minimization problem.

If the function $f(\mathbf{x}, \mathbf{y})$ is convex with respect to $\mathbf{x}$ and the set $\mathscr{X}$ is convex, then problem (2.50) - (2.56) is a convex program. Furthermore, if the function $f(\mathbf{x}, \mathbf{y})$ is linear with respect to $\mathbf{x}$ and the set $\mathscr{X}$ is convex polyhedron, then the problem is a second-order cone program.

3. Robust Portfolio Selection

In this chapter, we formulate a portfolio management problem by utilizing the risk measures presented in this thesis. We show that the upper bounds described in the first chapter lead to robust portfolio optimization and discuss several robust downside risk models. Moreover, we present an explicit formula for solving the robust portfolio problem with CVaR or VaR as a risk measure. We specify the constraint set $\mathscr{X}$ in the case of CVaR and WCVaR as a risk measure, which usually includes for example the requirement of minimum expected return that occures in the worst-case.

Suppose that an investor in the financial market can choose from n risky assets, or we can simly say, that the investor can decide how to divide his/her available capital between n assets. The vector $\mathbf{x} = (x_1, \ldots, x_n)^T \in \mathbb{R}^n$ represents the amount of the investments in the assets decided by the investor and the random vector $\mathbf{y} = (y_1, \ldots, y_n)^T \in \mathbb{R}^n$ represents the uncertain returns of the assets. Therefore, the total portfolio return is defined as $\mathbf{x}^T\mathbf{y}$. Naturally, the total portfolio loss is the negative of the return and thus the loss function $f(\mathbf{x}, \mathbf{y})$ is equal to $-\mathbf{x}^T\mathbf{y}$.

Identifying investment goals of the investor is the first task in constructing a portfolio. Important factor to take into account is his/her risk tolerance, i.e., his/her will to risk some money for the possibility of greater returns. Obviously, the possibility of greater returns comes at the expense of greater risk of losses, which is a principle known as the risk-return trade-off.

Important items to consider in portfolio construction are amount of capital to invest (initial wealth) and future capital needs. Suppose that the investor has an initial wealth w_0. In this case, it holds

$$\mathbf{e}^T\mathbf{x} = w_0. \tag{3.1}$$

In order to ensure diversification and satisfy the regulations, we include the bound constraints

$$\underline{\boldsymbol{x}} \leq \boldsymbol{x} \leq \bar{\boldsymbol{x}}, \tag{3.2}$$

where $\underline{\boldsymbol{x}}$ and $\bar{\boldsymbol{x}}$ are the given lower and upper bounds on the portfolios.

Let r_0 be the minimum expected return that the investor requires in the worst-case. This expectation can be represented as

$$\min_{p(\cdot)\in\mathscr{P}} \mathsf{E}_p(\boldsymbol{x}^T\boldsymbol{y}) \geq r_0, \tag{3.3}$$

where E_p denotes the expectation operator with respect to the distribution $p(\cdot)$ of $\mathbf{y}$.

In general, $\mathscr{X}$ is specified by (3.1), (3.2) and (3.3), i.e.,

$$\mathscr{X} = \{\boldsymbol{x} : \mathbf{e}^T\mathbf{x} = w_0, \underline{\boldsymbol{x}} \leq \boldsymbol{x} \leq \bar{\boldsymbol{x}}, \min_{p(\cdot)\in\mathscr{P}} \mathsf{E}_p(\boldsymbol{x}^T\boldsymbol{y}) \geq r_0\}. \tag{3.4}$$

3.1 Portfolio Selection with CVaR

In this case, condition (3.3) can be written as

$$\mathsf{E}_p(\boldsymbol{x}^T\boldsymbol{y}) \geq r_0, \tag{3.5}$$

where $p(\cdot)$ denotes the distribution of $\boldsymbol{y}$. We formulate this constraint as

$$\boldsymbol{x}^T\bar{\boldsymbol{y}} \geq r_0, \tag{3.6}$$

where $\bar{\boldsymbol{y}}$ denotes the expected value of $\boldsymbol{y}$ with respect to $p(\cdot)$.

By (2.4)- (2.7), the portfolio selection problem based on CVaR can be formulated as the linear program with decision variables $(\mathbf{x}, \mathbf{u}, \zeta) \in \mathbb{R}^n \times \mathbb{R}^S \times \mathbb{R}$:

$$\min\{2.4 : (2.6) - (2.7), (3.1), (3.2) \text{ and } (3.6)\}, \tag{3.7}$$

where $f(\mathbf{x}, \mathbf{y}_{(k)})$ in (2.6) is specified as $-\mathbf{x}^T\mathbf{y}_{(k)}$.

3.2 Robust Portfolio Selection Using LPM

In this section, we discuss the robust downside risk models. We are particularly interested in the lower partial moments $\mathrm{LPM}_m(r)$, $m \in \{0, 1, 2\}$.

The portfolio selection models (P_m) are formulated as

$$\begin{array}{lll} (P_m) & \min\limits_{\boldsymbol{x}} & \mathsf{E}\,[(r - \boldsymbol{x}^T\boldsymbol{y})_+^m], \\ & \text{s.t.} & \boldsymbol{x}^T\boldsymbol{e} = 1, \end{array}$$

where $m \in \{0, 1, 2\}$ represents risk aversion of the investor, and $\boldsymbol{e}$ denotes the vector of all ones with an appropriate dimension.

We assume that the mean and the covariance of $\mathbf{y}$ are known, and we make no further assumption on its distribution. For a given portfolio $\mathbf{x}$, our risk measure is $\mathsf{E}\,[(r - \boldsymbol{x}^T\boldsymbol{y})_+^m]$. The corresponding robust portfolio selection model may be written as

$$\begin{array}{rll} RP_m = & \min\limits_{\boldsymbol{x}} & \sup\limits_{\boldsymbol{y} \sim (\boldsymbol{\mu}, \Gamma)} \mathsf{E}\,[(r - \boldsymbol{x}^T\boldsymbol{y})_+^m] \\ & \text{s.t.} & \boldsymbol{x}^T\boldsymbol{e} = 1, \end{array}$$

where $m \in \{0, 1, 2\}$.

The upper bounds presented in the first chapter can be used to get an explicit expression for the objective function in RP_m. Those bounds are good only for univariate distributions, whereas the ambiguous distribution set P in RP_m involves multidimensional distributions. For any given $\mathbf{y}$, if we know the moments of $\mathbf{y}$, then we have all the information about the moments for the distribution $\mathbf{x}^T\mathbf{y}$. In general, the opposit is not true, but if we work only with the first two moments, then there is actually no loss of information.

For any $a \neq 0 \in \mathbb{R}^n$, we consider two sets

$$\begin{aligned} A &:= \{\boldsymbol{a}^T\boldsymbol{\xi} \,|\, \mathsf{E}\,[\boldsymbol{\xi}] = \boldsymbol{\mu}, \operatorname{cov}(\boldsymbol{\xi}) = \Gamma\}, \\ B &:= \{\eta \,|\, \mathsf{E}\,[\eta] = \boldsymbol{a}^T\boldsymbol{\mu}, \operatorname{var}(\eta) = \boldsymbol{a}^T\Gamma\boldsymbol{a}\}. \end{aligned}$$

Obviously, $A \subseteq B$. As proved in [6], the opposite relationship also holds. It means

$$A = B. \tag{3.8}$$

By this equivalence, we have

$$\sup_{\boldsymbol{y} \sim (\boldsymbol{\mu}, \Gamma)} \mathsf{E}\left[(r - \boldsymbol{x}^T \boldsymbol{y})_+^m\right] = \sup_{\zeta \sim (\boldsymbol{x}^T \boldsymbol{\mu}, \boldsymbol{x}^T \Gamma \boldsymbol{x})} \mathsf{E}\left[(r - \zeta)_+^m\right].$$

Therefore, we can directly apply the univariate bounds developed in the first chapter.

3.2.1 Explicit Solution of the Robust Portfolio Problem

In the following theorem, we present how to derive explicit solution of the robust portfolio problem based on LPM for the cases $m \in \{0, 1, 2\}$ (see [6]). Suppose $\Gamma \succ 0$. We denote

$$\begin{aligned} c_0 &:= \boldsymbol{e}^T \Gamma^{-1} \boldsymbol{e}, \quad c_1 := \boldsymbol{e}^T \Gamma^{-1} \boldsymbol{\mu}, \quad c_2 := \boldsymbol{\mu}^T \Gamma^{-1} \boldsymbol{\mu}, \\ b_0 &:= \frac{c_0}{c_0 c_2 - c_1^2}, \quad b_1 := \frac{c_1}{c_0 c_2 - c_1^2}, \quad b_2 := \frac{c_2}{c_0 c_2 - c_1^2}. \end{aligned}$$

Theorem 11 ([6])**.** *Consider the optimization problem*

$$\begin{aligned} v(RP_m) \quad &:= \quad \min_{\boldsymbol{x}} \sup_{\zeta \sim (\boldsymbol{x}^T \boldsymbol{\mu}, \boldsymbol{x}^T \Gamma \boldsymbol{x})} E\left[(r - \zeta)_+^m\right] \\ s.t. \quad & \quad \boldsymbol{x}^T \boldsymbol{e} = 1. \end{aligned}$$

For the case $m \in \{0, 1, 2\}$, we have the following explicit solutions:

(a) If $b_1 \geq r b_0$, then

$$\begin{aligned} v(RP_0) &= \frac{1}{1 + (\boldsymbol{\mu} - r\boldsymbol{e})^T \Gamma^{-1} (\boldsymbol{\mu} - r\boldsymbol{e})}, \\ \boldsymbol{x}^*_{RP_0} &= \frac{\Gamma^{-1}(\boldsymbol{\mu} - r\boldsymbol{e})}{\boldsymbol{e}^T \Gamma^{-1} \boldsymbol{\mu} - r \boldsymbol{e}^T \Gamma^{-1} \boldsymbol{e}}. \end{aligned}$$

Else, if $b_1 < r b_0$, then $v(RP_0) = 1/(1 + 1/b_0)$.

(b)

$$v(RP_1) = \frac{b_0 r - b_1 + \sqrt{(b_0 r - b_1)^2 + (b_0 b_2 - b_1^2)(b_0 + 1)}}{2(b_0 + 1)},$$

$$\boldsymbol{x}^*_{RP_1} = \begin{pmatrix} \Gamma^{-1}\boldsymbol{\mu} & \Gamma^{-1}\boldsymbol{e} \end{pmatrix} \begin{pmatrix} b_0 & -b_1 \\ -b_1 & b_2 \end{pmatrix} \begin{pmatrix} \frac{b_0(b_1 + r) + \sqrt{(b_0 r - b_1)^2 + (b_0 b_2 - b_1^2)(b_0 + 1)}}{b_0(b_0 + 1)} \\ 1 \end{pmatrix}.$$

(c)

$$\begin{aligned} v(RP_2) &= \frac{[(b_0 r - b_1)_+]^2}{b_0(b_0 + 1)} + \frac{1}{c_0}, \\ \boldsymbol{x}^*_{RP_2} &= \begin{pmatrix} \Gamma^{-1}\boldsymbol{\mu} & \Gamma^{-1}\boldsymbol{e} \end{pmatrix} \begin{pmatrix} b_0 & -b_1 \\ -b_1 & b_2 \end{pmatrix} \begin{pmatrix} \frac{(b_0 r - b_1)_+}{b_0(b_0 + 1)} + \frac{b_1}{b_0} \\ 1 \end{pmatrix}. \end{aligned}$$

Proof. Using the bounds presented in the first chapter, we denote

$$\begin{aligned} f_0(s,t) &:= \sup_{X\sim(s,t^2)} \mathsf{E}\left[(r-X)_+^0\right] = \frac{1}{1+(r-s)^2/t^2}; \\ f_1(s,t) &:= \sup_{X\sim(s,t^2)} \mathsf{E}\left[(r-X)_+^1\right] = \frac{r-s+\sqrt{t^2+(r-s)^2}}{2}, \\ f_2(s,t) &:= \sup_{X\sim(s,t^2)} \mathsf{E}\left[(r-X)_+^2\right] = [(r-s)_+]^2 + t^2. \end{aligned}$$

Now we can express our optimization problem as follows

$$\begin{aligned} \upsilon(RP_m) &= \min_{\boldsymbol{x}}\{f_m(\boldsymbol{x}^T\boldsymbol{\mu}, \sqrt{\boldsymbol{x}^T\Gamma\boldsymbol{x}})|\boldsymbol{x}^T\boldsymbol{e}=1\} \\ &= \min_{s\in\mathbb{R}}\min_{\boldsymbol{x}}\{f_m(s,\sqrt{\boldsymbol{x}^T\Gamma\boldsymbol{x}})|\boldsymbol{x}^T\boldsymbol{e}=1, \boldsymbol{x}^T\boldsymbol{\mu}=s\} \end{aligned} \tag{3.9}$$

For any given s and m, the optimal solution $\mathbf{x}_s^*$ of the inner optimization problem in (3.9) is a mean-variance efficient solution:

$$\begin{aligned} \boldsymbol{x}_s^* &= \arg\min_{\boldsymbol{x}}\{f_m(s,\sqrt{\boldsymbol{x}^T\Gamma\boldsymbol{x}})|\boldsymbol{x}^T\boldsymbol{e}=1, \boldsymbol{x}^T\boldsymbol{\mu}=s\} \\ &= \arg\min_{\boldsymbol{x}}\{\boldsymbol{x}^T\Gamma\boldsymbol{x}|\boldsymbol{x}^T\boldsymbol{e}=1, \boldsymbol{x}^T\boldsymbol{\mu}=s\} \\ &= \begin{pmatrix}\Gamma^{-1}\boldsymbol{\mu} & \Gamma^{-1}\boldsymbol{e}\end{pmatrix}\begin{pmatrix} b_0 & -b_1 \\ -b_1 & b_2\end{pmatrix}\begin{pmatrix} s \\ 1\end{pmatrix}. \end{aligned} \tag{3.10}$$

The second equality comes from the increasing property of $f_m(s,t)$ in t for all $m \in \{0,1,2\}$. It holds

$$(x_s^*)^T\Gamma x_s^* = b_0s^2 - 2b_1s + b_2, \tag{3.11}$$

and thus we have

$$\upsilon(RP_m) = \min_{s\in\mathbb{R}} f_m(s, \sqrt{b_0s^2-2b_1s+b_2}).$$

By solving the above problem we get

$$\begin{aligned} s^*_{RP_0} &= \arg\min_{s\in\mathbb{R}} f_0(s,\sqrt{b_0s^2-2b_1s+b_2}) \\ &= \arg\max_{s\geq r} \frac{(r-s)^2}{b_0s^2-2b_1s+b_2} \\ &= \begin{cases} \frac{b_2-b_1r}{b_1-b_0r}, & \text{if } b_1 \geq rb_0 \\ +\infty, & \text{if } b_1 \leq rb_0 \end{cases} \\ s^*_{RP_1} &= \arg\min_{s\in\mathbb{R}} f_1(s,\sqrt{b_0s^2-2b_1s+b_2}) \\ &= \arg\min_{s\in\mathbb{R}} \frac{r-s+\sqrt{b_0s^2-2b_1s+b_2+(r-s)^2}}{2} \\ &= \frac{b_0(b_1+r)+\sqrt{(b_0r-b_1)^2+(b_0b_2-b_1^2)(b_0+1)}}{b_0(b_0+1)} \end{aligned}$$

$$\begin{aligned} s^*_{RP_2} &= \arg\min_{s\in\mathbb{R}} f_2(s, \sqrt{b_0 s^2 - 2b_1 s + b_2}) \\ &= \arg\min_{s\in\mathbb{R}} [(r-s)_+]^2 + b_0 s^2 - 2b_1 s + b_2 \\ &= \frac{(b_0 r - b_1)_+}{b_0(b_0+1)} + \frac{b_1}{b_0}. \end{aligned}$$

This completes the proof. □

Remark. All of the above portfolios are mean-variance efficient. In particular, $s^*_{RP_0} \geq s^*_{RP_1} \geq s^*_{RP_2}$, which means that for a fixed r, the higher the order of lower partial moments, the more conservative portfolio.

With the simple constraint $\boldsymbol{x}^T\boldsymbol{e} = 1$, the robust portfolio problem with CVaR or VaR as risk measure can be solved explicitly.

Theorem 12 ([6])**.** *Suppose the loss function is $f(\boldsymbol{x}, \boldsymbol{y}) = -\boldsymbol{x}^T\boldsymbol{y}$ and the random vector $\boldsymbol{y}$ has mean $\boldsymbol{\mu}$ and covariance matrix $\Gamma \succ 0$. Let $\alpha \in (0.5, 1]$. Consider*

$$\begin{aligned} \upsilon(RC_\alpha) = \quad & \min_{\boldsymbol{x}} \quad \sup_{p\in P} CVaR_\alpha(\boldsymbol{x}) \\ & s.t. \quad \boldsymbol{x}^T\boldsymbol{e} = 1, \end{aligned}$$

and

$$\begin{aligned} \upsilon(RV_\alpha) = \quad & \min_{\boldsymbol{x}} \quad \sup_{p\in P} VaR_\alpha(\boldsymbol{x}) \\ & s.t. \quad \boldsymbol{x}^T\boldsymbol{e} = 1, \end{aligned}$$

Then the optimal value of worst-case CVaR is:

$$\upsilon(RC_\alpha) = \begin{cases} \frac{\sqrt{b_0 b_2 - b_1^2}\sqrt{\frac{\alpha b_0}{1-\alpha} - 1}}{b_0}, & \text{if } \frac{\alpha}{1-\alpha} b_0 \geq 1, \\ -\infty, & \text{if } \frac{\alpha}{1-\alpha} b_0 < 1, \end{cases}$$

and when $\alpha/(1-\alpha)b_0 \geq 1$, with optimal solution

$$\boldsymbol{x}^*_{RC_\alpha} = \begin{pmatrix} \Gamma^{-1}\boldsymbol{\mu} & \Gamma^{-1}\boldsymbol{e} \end{pmatrix} \begin{pmatrix} b_0 & -b_1 \\ -b_1 & b_2 \end{pmatrix} \begin{pmatrix} \frac{\sqrt{b_0 b_2 - b_1^2}}{b_0\sqrt{\alpha b_0/(1-\alpha) - 1}} + \frac{b_1}{b_0} \\ 1 \end{pmatrix}.$$

The optimal value of worst-case VaR is:

$$\upsilon(RV_\alpha) = \begin{cases} \frac{\sqrt{b_0 b_2 - b_1^2}\sqrt{b_0/(4\alpha(1-\alpha)) - b_0 - 1}}{b_0} - \frac{b_1}{b_0}, & \text{if } \frac{b_0}{4\alpha(1-\alpha)} \geq 1 + b_0, \\ -\infty, & \text{if } \frac{b_0}{4\alpha(1-\alpha)} < 1 + b_0, \end{cases}$$

When $b_0/(4\alpha(1-\alpha)) \geq 1 + b_0$, the optimal solution is

$$\boldsymbol{x}^*_{RV_\alpha} = \begin{pmatrix} \Gamma^{-1}\boldsymbol{\mu} & \Gamma^{-1}\boldsymbol{e} \end{pmatrix} \begin{pmatrix} b_0 & -b_1 \\ -b_1 & b_2 \end{pmatrix} \begin{pmatrix} \frac{\sqrt{b_0 b_2 - b_1^2}}{b_0\sqrt{b_0/(4\alpha(1-\alpha)) - b_0 - 1}} + \frac{b_1}{b_0} \\ 1 \end{pmatrix}.$$

The proof of the previous theorem, is given in [6]. In the following, we outline the main idea of this proof. By (2.1), for the linear loss function $f(\boldsymbol{x}, \boldsymbol{y}) = -\boldsymbol{x}^T\boldsymbol{y}$, the calculation of CVaR can be achieved by minimizing of the following auxiliary function

$$F_\alpha(\boldsymbol{x}, \zeta) = \zeta + \frac{1}{1-\alpha}\,\mathsf{E}\,[-\boldsymbol{x}^T\boldsymbol{y} - \zeta]_+,$$

where $[t]_+ = \max\{t, 0\}$. Using Lemma 3 and the proposed equivalence (3.8), we have

$$\begin{aligned}
\mathrm{RC}_\alpha(\boldsymbol{x}) &= \sup_{p\in D} \mathrm{CVaR}_\alpha(\boldsymbol{x}) \\
&= \min_{\zeta\in\mathbb{R}} \sup_{\xi\sim(\boldsymbol{x}^T\boldsymbol{\mu}, \boldsymbol{x}^T\Gamma\boldsymbol{x})} \zeta + \frac{1}{1-\alpha}\,\mathsf{E}\,[(-\zeta-\xi)_+] \\
&= \min_{\zeta\in\mathbb{R}} \zeta + \frac{1}{1-\alpha} \sup_{\xi\sim(\boldsymbol{x}^T\boldsymbol{\mu}, \boldsymbol{x}^T\Gamma\boldsymbol{x})} \mathsf{E}\,[(-\zeta-\xi)_+] \\
&= \min_{\zeta\in\mathbb{R}} \zeta + \frac{1}{1-\alpha}\left[\frac{1}{2}\sqrt{\boldsymbol{x}^T\Gamma\boldsymbol{x} + (\boldsymbol{x}^T\boldsymbol{\mu} + \zeta)^2} + \frac{-\zeta - \boldsymbol{x}^T\boldsymbol{\mu}}{2}\right]
\end{aligned}$$

Let

$$h_\alpha(\zeta) = \zeta + \frac{1}{1-\alpha}\left[\frac{1}{2}\sqrt{\boldsymbol{x}^T\Gamma\boldsymbol{x} + (\boldsymbol{x}^T\boldsymbol{\mu} + \zeta)^2} + \frac{-\zeta - \boldsymbol{x}^T\boldsymbol{\mu}}{2}\right],$$

and

$$\mathrm{RC}_\alpha(\boldsymbol{x}) = h(\zeta_x^*),$$

where ζ_x^* denotes the optimal solution minimizing $h_\alpha(\zeta)$ for given $\boldsymbol{x}$. By the first order optimality condition, (3.10) and (3.11), we get the required explicit expression.

Remark. As Zhu and Fukushima notice in [6], the expected downfall measured by CVaR and VaR are unbounded if $\alpha/(1-\alpha)b_0 < 1$ and $b_0/(4\alpha(1-\alpha)) \leq 1 + b_0$. In all other cases, the optimal portfolios are mean-variance efficient.

3.3 Robust Portfolio Selection Using WCVaR

In this section, we consider that random returns of financial assets are just specified by a set of distributions. We focus on robust portfolio selection problems under the three types of uncertainties described in the previous chapter.

The robust portfolio selection problem using WCVaR can be represented as

$$\min_{\mathbf{x}\in\mathscr{X}} \mathrm{WCVaR}_\alpha(\mathbf{x}).$$

According to the minimization problems that we formulated in the previous chapter, to complete the robust portfolio selection model, we only need to specify the constraint set $\mathscr{X}$.

3.3.1 Mixture Distribution Uncertainty

In the case of mixture distribution uncertainty given by (2.8), condition (3.3) can be written as

$$\sum_{i=1}^{l} \lambda_i \, \mathsf{E}_{p^i}(\boldsymbol{x}^T \boldsymbol{y}) \geq r_0, \forall \boldsymbol{\lambda} \in \Lambda, \tag{3.12}$$

where Λ is defined by (2.9). It is easy to see that the constraint (3.12) is equivalent to

$$\mathsf{E}_{p^i}(\boldsymbol{x}^T \boldsymbol{y}) \geq r_0, i = 1, \ldots, l. \tag{3.13}$$

We represent the set of contraints (3.13) as

$$\boldsymbol{x}^T \bar{\boldsymbol{y}}^i \geq r_0, i = 1, \ldots, l, \tag{3.14}$$

where $\bar{\boldsymbol{y}}^i$ denotes the expected value of $\mathbf{y}$ with respect to the likelihood distribution $p^i(\cdot)$. By (2.25)- (2.29), the robust portfolio selection problem under the mixture distribution is formulated as the following linear program with decision variables $(\mathbf{x}, \mathbf{u}, \zeta, \theta) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R} \times \mathbb{R}$:

$$\min\{\theta : (2.27) - (2.29), (3.1), (3.2) \text{ and } (3.14)\}, \tag{3.15}$$

where $f(\mathbf{x}, \mathbf{y}^i_{(k)})$ in (2.28) is specified as $-\mathbf{x}^T \mathbf{y}^i_{(k)}$.

3.3.2 Box Uncertainty in Discrete Distribution

The sample space of random vector $\mathbf{y}$ is given by $\{\mathbf{y}_{(1)}, \mathbf{y}_{(2)}, \ldots, \mathbf{y}_{(S)}\}$. We denote

$$Y = \begin{bmatrix} \boldsymbol{y}^T_{(1)} \\ \vdots \\ \boldsymbol{y}^T_{(S)} \end{bmatrix}. \tag{3.16}$$

In the case of box uncertainty in discrete distributions, by (2.35) and (3.4), the constraint set $\mathscr{X}$ is specified by

$$\mathscr{X}_B = \{\mathbf{x} : \mathbf{e}^T \mathbf{x} = w_0, \underline{\boldsymbol{x}} \leq \boldsymbol{x} \leq \bar{\boldsymbol{x}}, (Y\mathbf{x})^T \boldsymbol{\pi}^0 + \min_{\{\boldsymbol{\eta} : \mathbf{e}^T \boldsymbol{\eta} = 0, \underline{\boldsymbol{\eta}} \leq \boldsymbol{\eta} \leq \bar{\boldsymbol{\eta}}\}} (Y\mathbf{x})^T \boldsymbol{\eta} \geq \mu\}.$$

The dual problem of the inner linear program

$$\min_{\boldsymbol{\eta} \in \mathbb{R}^S} \{(Y\mathbf{x})^T \boldsymbol{\eta} : \mathbf{e}^T \boldsymbol{\eta} = 0, \underline{\boldsymbol{\eta}} \leq \boldsymbol{\eta} \leq \bar{\boldsymbol{\eta}}\}$$

is formulated as

$$\max_{(\delta, \boldsymbol{\tau}, \boldsymbol{\nu}) \in \mathbb{R} \times \mathbb{R}^S \times \mathbb{R}^S} \{\bar{\boldsymbol{\eta}}^T \boldsymbol{\tau} + \underline{\boldsymbol{\eta}}^T \boldsymbol{\nu} : \boldsymbol{e}\delta + \boldsymbol{\tau} + \boldsymbol{\nu} = Y\boldsymbol{x}, \boldsymbol{\tau} \leq 0, \boldsymbol{\nu} \geq 0\}.$$

Define

$$\Phi^B = \{(\boldsymbol{x}, \delta, \boldsymbol{\tau}, \boldsymbol{\nu}) : \begin{array}{ll} \boldsymbol{e}^T\boldsymbol{x} & = w_0, \underline{\boldsymbol{x}} \leq \boldsymbol{x} \leq \bar{\boldsymbol{x}}, \boldsymbol{e}\delta + \boldsymbol{\tau} + \boldsymbol{\nu} = Y\boldsymbol{x}, \\ \boldsymbol{\tau} & \leq 0, \boldsymbol{\nu} \geq 0, (Y\boldsymbol{x})^T\boldsymbol{\pi}^0 + \bar{\boldsymbol{\eta}}^T\boldsymbol{\tau} + \underline{\boldsymbol{\eta}}^T\boldsymbol{\nu} \geq \mu\} \end{array}$$

and

$$\Phi^B_{\mathscr{X}} = \{\mathbf{x} : \exists(\delta, \boldsymbol{\tau}, \boldsymbol{\nu}) \text{ such that } (\mathbf{x}, \delta, \boldsymbol{\tau}, \boldsymbol{\nu}) \in \Phi^B\}.$$

By the duality theory of linear programming, it holds [17]

$$\mathscr{X}_B = \Phi^B_{\mathscr{X}}. \tag{3.17}$$

By (2.38) - (2.44) and (3.17), in the case of box uncertainty in discrete distribution, the robust portfolio selection problem can be formulated as the following linear program with decision variables $(\mathbf{x}, \mathbf{u}, z, \boldsymbol{\xi}, \boldsymbol{\omega}, \zeta, \theta, \delta, \boldsymbol{\tau}, \boldsymbol{\nu}) \in \mathbb{R}^n \times \mathbb{R}^S \times \mathbb{R} \times \mathbb{R}^S \times \mathbb{R}^S \times \mathbb{R} \times \mathbb{R} \times \mathbb{R} \times \mathbb{R}^S \times \mathbb{R}^S$:

$$\min\{\theta : (2.40) - (2.44) \text{ and } (\mathbf{x}, \delta, \boldsymbol{\tau}, \boldsymbol{\nu}) \in \Phi^B\}, \tag{3.18}$$

where $f(\mathbf{x}, \mathbf{y}_{(k)})$ is equal to $-\mathbf{x}^T\mathbf{y}_{(k)}$.

3.3.3 Ellipsoidal Uncertainty in Discrete Distribution

By (2.47) and (3.4), in the case of the ellipsoidal uncertainty in discrete distributions, $\mathscr{X}$ is specified by

$$\mathscr{X}_E = \{\mathbf{x} : \begin{array}{ll} \mathbf{e}^T\mathbf{x} & = w_0, \underline{\boldsymbol{x}} \leq \boldsymbol{x} \leq \bar{\boldsymbol{x}}, (Y\mathbf{x})^T\boldsymbol{\pi}^0 \\ + & \min\limits_{\{\boldsymbol{\eta}:\mathbf{e}^T A\boldsymbol{\eta}=0, \boldsymbol{\pi}^0 + A\boldsymbol{\eta} \geq 0, \|\boldsymbol{\eta}\| \leq 1\}} (Y\mathbf{x})^T A\boldsymbol{\eta} \geq \mu\}, \end{array}$$

where Y is definied by (3.16). Probably the most difficult part to calculate is the minimum operation in the constraints, which is formulated as the second-order cone program

$$\min_{\boldsymbol{\eta} \in \mathbb{R}^S} \{(Y\mathbf{x})^T A\boldsymbol{\eta} : \mathbf{e}^T\mathbf{x} = w_0, \underline{\boldsymbol{x}} \leq \boldsymbol{x} \leq \bar{\boldsymbol{x}}, \mathbf{e}^T A\boldsymbol{\eta} = 0, \boldsymbol{\pi}^0 + A\boldsymbol{\eta} \geq 0, \|\boldsymbol{\eta}\| \leq 1\}$$

and the dual program is given by [17]

$$\max_{(\sigma, \boldsymbol{\tau}, \boldsymbol{\nu}, \delta) \in \mathbb{R} \times \mathbb{R}^S \times \mathbb{R}^S \times \mathbb{R}} \{-\sigma - (\boldsymbol{\pi}^0)^T\boldsymbol{\tau} : \boldsymbol{\nu} + A^T\boldsymbol{\tau} + A^T\boldsymbol{e}\delta = A^T Y\boldsymbol{x}, \|\boldsymbol{\nu}\| \leq \sigma, \boldsymbol{\tau} \geq 0\}.$$

Define

$$\Phi^E = \{(\boldsymbol{x}, \delta, \boldsymbol{\tau}, \boldsymbol{\nu}, \sigma) : \begin{array}{rcl} \boldsymbol{e}^T\boldsymbol{x} & = & w_0, \underline{\boldsymbol{x}} \leq \boldsymbol{x} \leq \bar{\boldsymbol{x}}, \boldsymbol{e}\delta + \boldsymbol{\tau} + \boldsymbol{\nu}) = Y\boldsymbol{x}, \\ \boldsymbol{\nu} + A^T\boldsymbol{\tau} + A^T\boldsymbol{e}\delta & = & A^T Y\boldsymbol{x}, \|\boldsymbol{\nu}\| \leq \sigma, \\ \boldsymbol{\tau} \geq 0, (Y\boldsymbol{x})^T\boldsymbol{\pi}^0 & - & \sigma - (\boldsymbol{\pi}^0)^T\boldsymbol{\tau} \geq \mu\} \end{array}$$

and

$$\Phi^E_{\mathscr{X}} = \{\mathbf{x} : \exists(\delta, \boldsymbol{\tau}, \boldsymbol{\nu}, \sigma) \text{ such that } (\mathbf{x}, \delta, \boldsymbol{\tau}, \boldsymbol{\nu}, \sigma) \in \Phi^E\}.$$

By the conic duality theory [10], we have

$$\mathscr{X}_E = \Phi^E_{\mathscr{X}}. \tag{3.19}$$

By (2.50) - (2.56) and (3.19), the robust portfolio selection can be written as the following second-order cone program:

$$\min\{\theta : (2.52) - (2.56) \text{ and } (\mathbf{x}, \delta, \boldsymbol{\tau}, \boldsymbol{\nu}, \sigma) \in \Phi^E\}, \tag{3.20}$$

with decision variables $(\mathbf{x}, \mathbf{u}, \gamma, \boldsymbol{\omega}, \boldsymbol{\xi}, z, \zeta, \theta, \sigma, \boldsymbol{\tau}, \boldsymbol{\nu}, \delta) \in \mathbb{R}^n \times \mathbb{R}^S \times \mathbb{R} \times \mathbb{R}^S \times \mathbb{R}^S \times \mathbb{R} \times \mathbb{R} \times \mathbb{R} \times \mathbb{R} \times \mathbb{R}^S \times \mathbb{R}^S \times \mathbb{R}$. The function $f(\mathbf{x}, \mathbf{y}_{(k)})$ is specified as $-\mathbf{x}^T\mathbf{y}_{(k)}$.

4. Numerical Application

In this chapter, we present the results of numerical applications for nominal and robust portfolio optimization problems performed via the methods proposed in this thesis. Firstly, we apply the original CVaR approach and examine its performance. Secondly, we present the results for the robust portfolio optimization under box uncertainty in distribution and illustrate its advantage over the nominal formulation. Finally, we consider mixture distribution. We look at the benefits as well as the shortcomings of the robust portfolio optimization problem and present an explicit solution of the worst-case CVaR. As we will see later, the dataset used in the first two examples is not suitable for this numerical experiment, because daily returns seem to follow the same distribution within the whole time horizon. Therefore, in this example, we use another dataset where it makes sense to consider several time periods. Among other things, we also demonstrate a saying that goes "there is no such thing as a free lunch", which means that if an investor wants higher return he/she needs to take on lager risk.

Several important issues in the modeling, such as how to select stocks to be included in the portfolio or transaction costs, are beyond the scope of this thesis. However, the application and comparison of CVaR and several worst-case CVaR approaches, which is our main goal, is not undermine.

We use the historical market data (closing stock prices in USD) collected from Yahoo!Finance. All stocks that we work with are traded on New York Stock Exchange (NYSE).

The numerical applications are implemented on a PC (8.00 GB RAM, CPU 2.30 GHz) and all the problems are successfully solved within 5 seconds. We use a software environment **RStudio** for the computations and graphics. For solving large-scale linear programmimg we use Rglpk package (CPLEX solver).

4.1 CVaR Approach

In this example, we consider the financial assets of ten different companies, that are known as the companies with the top most expensive stocks as of January 2016, to construct the portfolio.

- Amazon.com, Inc. (AMZN);
- AutoZone, Inc. (AZO);
- Berkshire Hathaway Inc. (BRK-B);
- Chipotle Mexican Grill, Inc. (CMG);
- Alphabet Inc. (GOOG);
- Intuitive Surgical, Inc. (ISRG);
- Markel Corp. (MKL);
- NVR, Inc. (NVR);

Table 4.1: Estimated means, variances, and skewness of returns

	AMZN	AZO	BRK-B	CMG	GOOG
Mean (10^{-2})	0.1247	0.0874	0.0462	0.0810	0.0858
Variance (10^{-2})	0.0428	0.0129	0.0136	0.0401	0.0256
Skewness	0.5132	0.1135	0.6298	-0.5592	1.5258
	ISRG	MKL	NVR	PCLN	SEB
Mean (10^{-2})	0.0759	0.0713	0.0802	0.1075	0.0559
Variance (10^{-2})	0.0393	0.0125	0.0209	0.0367	0.0526
Skewness	0.4262	-0.3550	-0.7914	-0.2748	0.2497

- The Priceline Group Inc. (PCLN);
- Seaboard Corp. (SEB).

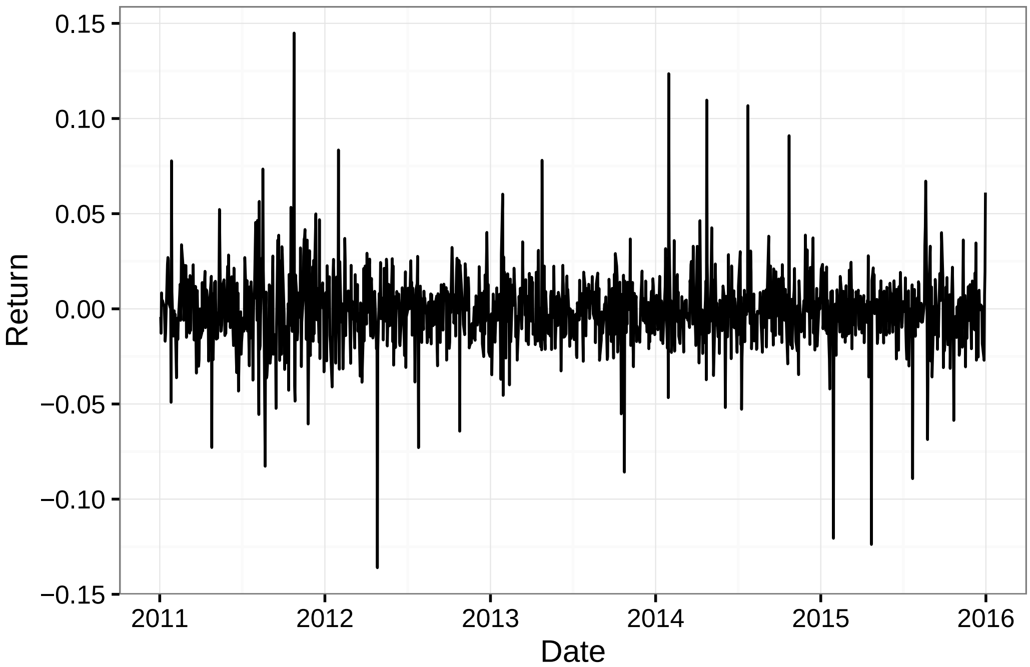

Figure 4.1: Daily returns of AMZN

A total of 1257 sample points of daily returns for these ten assets are collected from January 4, 2011 to December 30, 2015. Figure 4.1 is constructed by the sample points of daily returns of AMZN. The estimated means, variances, and skewness of returns are listed in Table 4.1.

In this example, we assume discrete distribution of random returns with realizations r_{ik}, $i = 1, \dots, 10$, $k = 1, \dots, 1257$, and probabilities 1/1257. We use linear programmimg model (3.7) for different values of minimal required return r_0 and confidence level α. We set $w_0 = 1$, $\underline{\boldsymbol{x}} = (0, \dots, 0)^T$ and $\bar{\boldsymbol{x}} = (1, \dots, 1)^T$.

Let r_{min} and r_{max} denote the minimum and maximum estimated means that are equal to $0.0462 \cdot 10^{-2}$ and $0.1247 \cdot 10^{-2}$, respectively. Figure 4.2 illustrates the values of optimal $\text{CVaR}_{0.95}$ for different values of the required minimal expected

Table 4.2: Comparison of CVaRs and expected returns for different values of r_0 and α

	CVaR_α			Expected return (10^{-2})		
r_0 (10^{-2})	$\alpha = 0.90$	$\alpha = 0.95$	$\alpha = 0.98$	$\alpha = 0.90$	$\alpha = 0.95$	$\alpha = 0.98$
0	0.0147	0.0189	0.0247	0.0773	0.0792	0.0736
0.05	0.0147	0.0189	0.0247	0.0773	0.0792	0.0736
0.075	0.0147	0.0189	0.0247	0.0773	0.0792	0.075
0.08	0.0147	0.0189	0.0250	0.08	0.08	0.08
0.085	0.0152	0.0192	0.0253	0.085	0.085	0.085
0.09	0.0163	0.0204	0.0262	0.09	0.09	0.09
0.095	0.0178	0.0224	0.0285	0.095	0.095	0.095
0.1	0.0196	0.0249	0.0321	0.1	0.1	0.1

return r_0 from interval $[r_{min}, r_{max}]$. Naturally, higher required minimal expected return leads to higher associated risk. Figure 4.2 demonstrates the increasing value of optimal $\text{CVaR}_{0.95}$ for the increasing value of r_0.

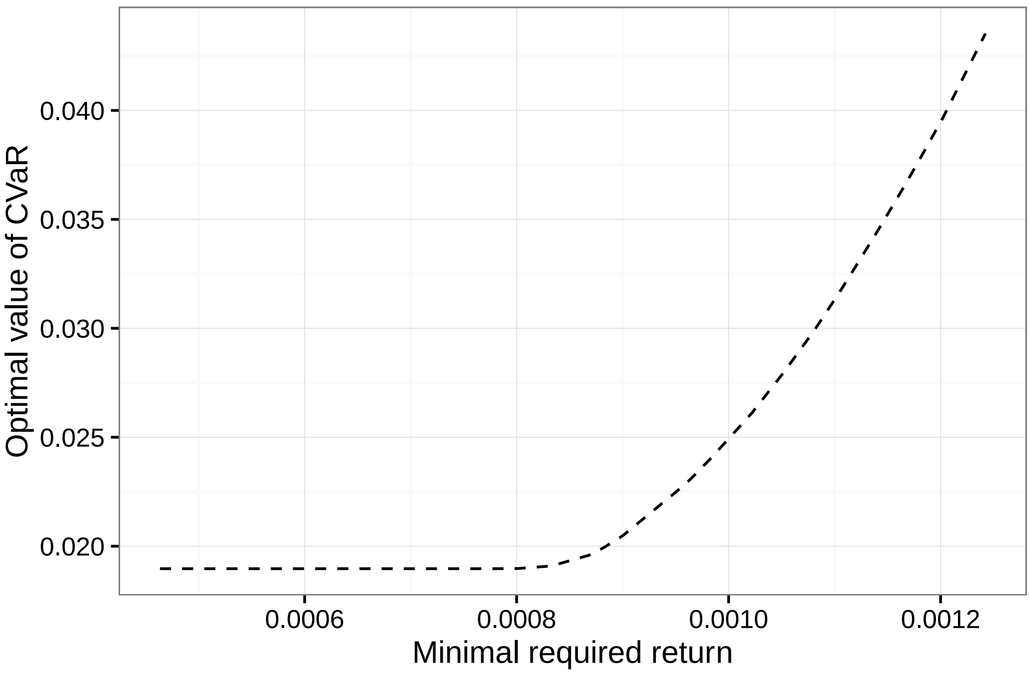

Figure 4.2: Optimal values of $\text{CVaR}_{0.95}$ for different values of minimal required return r_0

Table 4.2 shows the expected portfolio returns and the CVaRs of the corresponding optimal portfolios for different values of r_0 and α. If $\alpha \in \{0.09, 0.95\}$, the constraint $\mathbf{x}^T\bar{\mathbf{y}} > r_0$ is inactive for $r_0 \in \{0, 0.0005, 0.00075\}$ and active at an optimal solution for $r_0 \in \{0.0008, 0.00085, 0.0009, 0.00095, 0.001\}$. If $\alpha = 0.98$, the constraint on the minimal required return becomes active for $r_0 = 0.00075$. We can observe that the increasing value of α leads to higher values of CVaR_α. Obviously, if the minimal required return r_0 is higher than r_{max}, the optimization problem has no feasible solution.

4.2 WCVaR under Box Uncertainty

In this part, we perform market data simulation analysis for the robust portfolio optimization problem under the box uncertainty in distribution. As Zhu and Fukushima notice in their paper [17], a nonempty ellipsoid must contain a smaller box, and at the same time, must be contained by a bigger box. Thus, for both ellipsoidal and box uncertainties, we may expect that the simulation results will be very similar to each other. To avoid duplicate statements, we consider here only the case of box uncertainty, i.e., the linear programming model (3.18).

We use the dataset given in the previous example, where the portfolio is to be constructed by the ten assets. We set $\alpha = 0.95$, $w_0 = 1$, $\underline{\boldsymbol{x}} = (0, \ldots, 0)^T$ and $\bar{\boldsymbol{x}} = (1, \ldots, 1)^T$. The result of a robust portfolio problem depends on the structure of the uncertainty set. Thus, a properly chosen uncertainty set is the key to successful application. In this numerical experiment, the question is how to choose the parameters $\underline{\boldsymbol{\eta}}$ and $\bar{\boldsymbol{\eta}}$. In all computations, we set these paratemers symmetrically, i.e., $|\underline{\boldsymbol{\eta}}| = |\bar{\boldsymbol{\eta}}|$, and we denote their common absolute value as parameter $\boldsymbol{\eta}$.

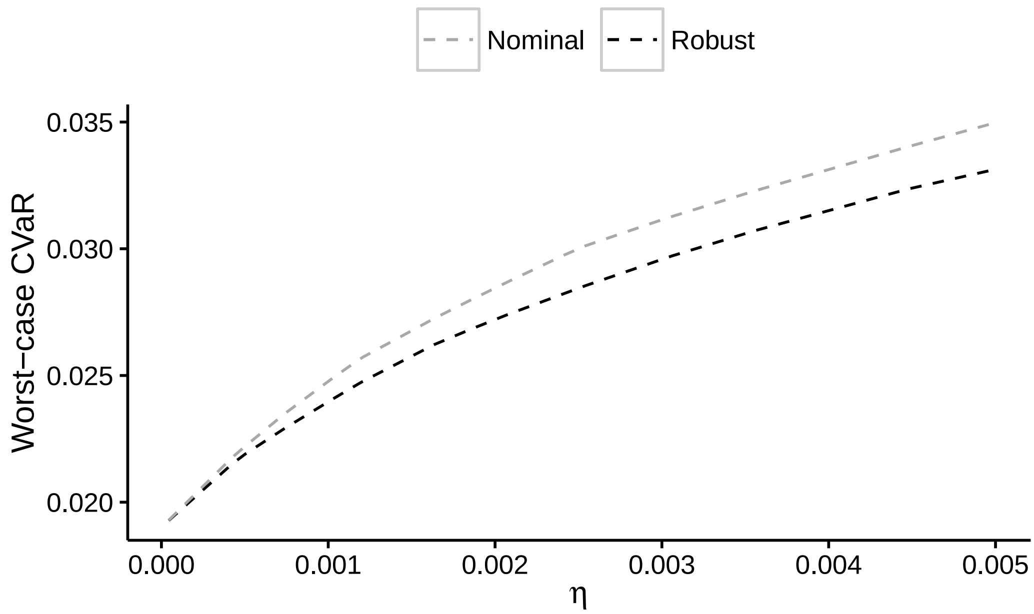

Figure 4.3: Worst-case CVaRs of nominal optimal and robust optimal portfolios.

The nominal optimal portfolio is obtained by solving model (3.18) with parameters $\underline{\boldsymbol{\eta}} = \bar{\boldsymbol{\eta}} = \mathbf{0}$. In this special case, the problem reduces to the original CVaR minimization problem as we mentioned in the previous chapter. The worst-case CVaR of the nominal optimal portfolio is obtained from model (3.18) by setting fixed values of $\boldsymbol{x} =$"nominal optimal portfolio".

For both nominal optimal and robust optimal portfolios, we compute the worst-case CVaR for different values of $\boldsymbol{\eta}$, and the fixed value of $r_0 = -0.05$. Figure 4.3, which illustrates a part of these numerical results, shows that the worst-case CVaR grows as the uncertainty grows. The most important observation is that the gap between the two curves in Figure 4.3 becomes larger as the

Table 4.3: Worst-case CVaR of nominal optimal and robust optimal portfolios for different values of $\boldsymbol{\eta}$ and r_0.

	$\boldsymbol{\eta}$					
	0.00001		0.00002		0.00003	
r_0 (10^{-2})	Robust	Nominal	Robust	Nominal	Robust	Nominal
0	0.019040	0.019004	0.019116	0.019114	0.019181	0.019188
0.05	0.019040	0.019004	0.019116	0.019114	0.019281	0.019188
0.075	0.019613	—	0.021850	—	—	—
0.08	0.020101	—	0.025927	—	—	—
0.085	0.022205	—	0.033132	—	—	—
0.09	0.025310	—	—	—	—	—
0.095	0.029380	—	—	—	—	—
0.1	0.034718	—	—	—	—	—

Table 4.4: Out-of-sample total return of nominal optimal and robust optimal portfolios for different values of $\boldsymbol{\eta}$ and r_0.

	$\boldsymbol{\eta}$				
	0	0.00001		0.00002	
r_0 (10^{-2})		Robust	Nominal	Robust	Nominal
0	0.09046	0.09051	0.09046	0.09154	0.09046
0.05	0.09046	0.09051	0.09046	0.09154	0.09046
0.075	0.09046	0.08427	—	0.09854	—
0.08	0.08593	0.09581	—	0.11056	—
0.085	0.09188	0.10280	—	0.11881	—
0.09	0.09013	0.11042	—	—	—
0.095	0.10147	0.11174	—	—	—
0.1	0.10999	0.12195	—	—	—

uncertainty grows, which demonstrates the advantage of the robust optimization formulation.

Table 4.3 shows the results corresponding to several values of r_0 and $\boldsymbol{\eta}$. Those results indicate that the portfolio problem becomes infeasible when r_0 or $\boldsymbol{\eta}$ increases to a certain value. For example, the nominal optimal portfolio obtained by solving (3.18) with $\boldsymbol{\eta} = 0$ and $r_0 = 0.00075$ is infeasible to (3.18) with $\boldsymbol{\eta} = 0.00001$, although problem (3.18) with $\boldsymbol{\eta} = 0.00001$ itself is feasible. For both nominal optimal and robust optimal portfolios, the values of the worst-case CVaRs for $\boldsymbol{\eta} = 0$ are equal to the values of $\text{CVaR}_{0.95}$ from the previous example. The corresponding values are listed in Table 4.2.

For out-of-sample analysis we use 124 sample points from January 4, 2016 to Jun 30, 2016. The total returns for different values of $\boldsymbol{\eta}$ and r_0 are listed in Table 4.4. We can observe that the robust approach leads to higher returns than nominal approach and CVaR approach ($\boldsymbol{\eta} = 0$).

Table 4.5: Estimated means, variances, and skewness of returns in different periods.

	Mean (10^{-2})		Variance (10^{-2})		Skewness	
	Period 1	Period 2	Period 1	Period 2	Period 1	Period 2
MGT	-0.0343	0.1486	0.0560	1.0291	1.4056	1.5693
ZION	-0.0393	0.4640	0.0220	1.1760	-0.0955	1.4979
MAC	0.1135	0.1716	0.0694	1.0251	11.2359	1.5782
MKL	0.0304	0.0383	0.0151	0.0935	0.0529	0.1863
MS	-0.0295	0.0965	0.0351	0.3363	-0.7099	4.3789

4.3 WCVaR under Mixture Distribution

The financial crisis 2008 had an impact on global stock markets, where securities suffered large losses during 2008 and early 2009. As we know, stock prices are volatile and can change extremely every day as a result of market forces. In this numerical experiment, we focus on companies that experienced several significant "ups and downs" in 2008 and 2009, and their stock prices have stayed unstable even after the active phase of the crisis. Nowadays, most of this companies are known as suitable for investors who prefer higher risk.

We consider the financial assets of the following five companies:

- MGT Capital Investments, Inc. (MGT);
- Zions Bancorporation (ZBK);
- The Macerich Company (MAC);
- Markel Corp. (MKL);
- Morgan Stanley (MS).

A total of 1600 sample points of daily returns for these assets are collected from January 4, 2005 to May 11, 2011. Figure 4.4 is constructed by the sample points of daily returns of MGT. We may observe that the distributions of returns vary before and after March 2008. The daily returns of the other four assets behave similarly. Due to this observation, we divide the whole time horizon into the following two subintervals, where each time period includes 800 sample points:

- Period 1: January 4, 2005 - March 6, 2008;
- Period 2: March 7, 2008 - May 9, 2011.

The estimated means, variances, and skewness of returns corresponding to different time periods are listed in Table 4.5. As we can see, the estimation of the statistical parameters is not stable. The Shapiro-Wilk test of normality rejected the null hypothesis that daily returns are normally distributed. Due to this fact we use a non-parametric test to detect significant differences in behaviour across the periods. More specifically, Friedman test shows that there exist remarkable difference in the distributions of returns between the two time periods for MGT, ZBK and MAC, whereas there is no significant difference for MKL and MS.

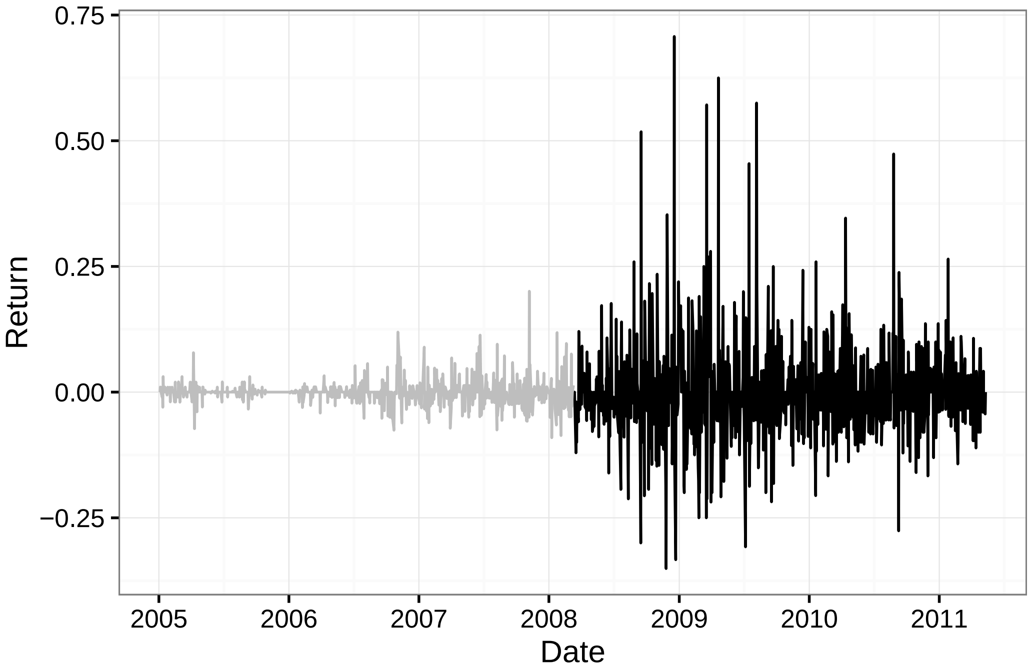

Figure 4.4: Daily returns of MGT

The original CVaR is usually used as a risk measure when the underlying assumption that the probability distribution is precisly known to be nominal one is met. In our example, where it is questionable to consider that all the samples are generated by an identical nominal probability distribution, it makes sense to assume a mixture distribution and perform the worst-case CVaR minimization.

According to the previous discussion, we consider the mixture distribution of two likelihood distributions and we assume that the sample points within each time period are generated by the corresponding likelihood distribution. We use setting $\alpha = 0.95$, $w_0 = 1$, $\underline{\boldsymbol{x}} = (0, 0, 0, 0, 0)^T$ and $\bar{\boldsymbol{x}} = (1, 1, 1, 1, 1)^T$.

In the special case, when $l = 1$, the problem (3.15) reduces to the original CVaR minimization problem. Thus, numerical experiments for the nominal portfolio optimization problem can be also performed via the linear programming model (3.7). In this numerical experiment, both nominal and robust portfolio optimization problems are performed via the model (3.15), where we set a corresponding parameter l. In the computation of the nominal portfolio optimization problem, we consider $l = 1$ and $S^1 = 1600$, i.e., all the sample points are used in the model and we assume that they are generated by one nominal probability distribution. In the computation of the robust portfolio optimization problem, we set $l = 2$ and $S^1 = S^2 = 800$, and we assume that the sample points within each time period are generated by the corresponding likelihood distribution.

We set different values of the minimal required return r_0. Table 4.6 shows the worst-case CVaRs at confidence level 0.95 of the corresponding optimal portfolios for each time period. The expected return of the corresponding optimal portfolios in different time periods are listed in Table 4.7.

We note that for the same value of r_0, the risk of the robust optimal portfolio optimization policy is larger than the risk of the nominal optimal portfolio policy.

Table 4.6: Comparison of associated risk of nominal optimal and robust optimal portfolios.

r_0 (10^{-2})	$\text{WCVaR}_{0.95}$ Nominal (l=1)	Robust (l=2)
0	0.0469	0.0614
0.05	0.0469	0.0687
0.06	0.0470	0.0794
0.07	0.0484	0.0986
0.08	0.0509	0.1197
0.09	0.0542	0.1426
0.1	0.0588	0.1664
0.15	0.0902	—
0.2	0.1383	—

Table 4.7: Comparison of expected return of nominal optimal and robust optimal portfolios.

r_0 (10^{-2})	Expected return (10^{-2})			
	Nominal		Robust	
	Period 1	Period 2	Period 1	Period 2
0	0.0236	0.0897	0.0248	0.0983
0.05	0.0236	0.0897	0.05	0.0708
0.06	0.0236	0.0964	0.06	0.0833
0.07	0.0246	0.1154	0.07	0.0990
0.08	0.0269	0.1331	0.08	0.1148
0.09	0.0266	0.1534	0.09	0.1305
0.10	0.0250	0.1750	0.10	0.1462
0.15	0.0159	0.2840	—	—
0.20	-0.0139	0.4137	—	—

The difference between associated risk of nominal optimal and robust optimal portfolios becomes more significant with increasing value of r_0.

In our example, the constraint $\mathsf{E}_{\text{nominal}}(\mathbf{x}^T\mathbf{y}) \geq r_0$ in the nominal portfolio optimization problem is inactive at an optimal solution for $r_0 \in \{0, 0.0005\}$, and it becomes active for $r_0 \in \{0.0006, 0.0007, 0.0008, 0.0009, 0.001, 0.0015, 0.002\}$. In the robust portfolio optimization problem, the constraint $\min\limits_{p(\cdot)\in\mathscr{P}} \mathsf{E}_p(\mathbf{x}^T\mathbf{y}) \geq r_0$ is inactive for each selected r_0. For $r_0 \in \{0.0015, 0.002\}$, the robust portfolio optimization problem is infeasible. According to the constraints (3.14), i.e., $\boldsymbol{x}^T\bar{\boldsymbol{y}}^i \geq r_0, i = 1, \ldots, l$, the robust optimal portfolio policy always guarantees the minimal required return r_0 for each time period. As the maximum in the Period 1 is equal to 0.001135, the condition (3.14) for $l = 2$ and $r_0 \in \{0.0015, 0.002\}$ is violated for an arbitrary portfolio. This results in infeasibility of robust portfolio optimization problem for any $r_0 > 0.001135$.

The means and variances computed by the total 1600 sample points for different values of r_0 are listed in Table 4.8. In any case, the mean and variance of the robust optimal portfolios are both larger than those of the nominal optimal portfolios. This observation suggests that robust optimal portfolios can be

Table 4.8: Mean and variance of nominal optimal and robust optimal portfolios (1600 sample points)

	Nominal portfolio		Robust portfolio	
r_0 (10^{-2})	Mean (10^{-2})	Variance (10^{-2})	Mean (10^{-2})	Variance (10^{-2})
0	0.0566	0.0436	0.0616	0.0443
0.05	0.0566	0.0436	0.0604	0.0617
0.06	0.06	0.0439	0.0717	0.0903
0.07	0.07	0.0467	0.0845	0.1374
0.08	0.08	0.0529	0.0974	0.2021
0.09	0.09	0.0620	0.1102	0.2844
0.1	0.10	0.0739	0.1231	0.3843

Table 4.9: Weights of the assets in the optimal portfolio

Asset	MGT	ZION	MAC	MKL	MS
Weight	0.0774	0.0602	0.0828	0.7855	-0.00598

more agressive than the nominal optimal portfolios for the same r_0. In general, an aggressive investment strategy emphasizes capital appreciation as a primary investment goal, rather than safety of capital.

Furthermore, we find that in the sense of worst-case trade-off, the nominal optimal policy generated by setting $r_0 = 0.002$ is dominated by the robust optimal policy generated by setting $r_0 = 0.0005$ because we have $0.0005 > -0.01390$ for the expected returns and $0.0687 < 0.1383$ for the associated risks. This fact together with Table 4.8 proposes that the worst-case requirement in the robust portfolio formulation does not influence the average performance of the portfolio substantially.

In this example, we also present an explicit solution of the robust portfolio model. The corresponding covariance matrix $\mathbf{\Gamma}$ is positive semidefinite, which means that the robust portfolio problem with $\mathrm{CVaR}_{0.95}$ as a risk measure can be solved explicitly. Using Theorem 12, the solution of $v(\mathrm{RC}_\alpha)$ is equal to 0.0899 and the expected return of the optimal portfolio in Period 1 and Period 2 is equal to 0.000284 and 0.00082, respectively. In comparison with the previous nominal approach, the value of $\mathrm{WCVaR}_{0.95}$ is significantly lower for the cases where the value of expected return for the whole time horizon is approximatelly equal to 0.00005.

In the case of the explicit approach, the weights of the optimal portfolio are listed in Table 4.9. As we can see, short-selling is allowed and used in this investment strategy.

Conclusion

This thesis studies different robust methods in portfolio theory. We introduced several robust approaches and formulated the corresponding portfolio optimization problems. We investigated the problem of minimizing the worst-case CVaR associated with mixture distribution uncertainty, box uncertainty, and ellipsoidal uncertainty in the distributions. In the case of mixture distribution, the proposed approximation produces a linear program. In the case of discrete distribution, we must deal with the max operation involved in the constraints. Using the linear and conic duality theory, under box uncertainty and ellipsoidal uncertainty in distributions, the problem can be cast as a linear and a second-order cone program, respectively. To complete the formulation of the robust portfolio selection model, we specified the corresponding constraint sets.

In the last chapter, we presented the results of numerical applications. Firstly, we applied the original CVaR approach. We showed that the higher required minimal expected return leads to higher associated risk. The associated risk also grows with the increasing value of the confidence level α.

Secondly, we considered the problem of minimizing the worst-case CVaR associated with box uncertainty in the distributions. We note again that a suitable specification of the uncertainty set is the key for successful practical application. We illustrated that the worst-case CVaR grows as the uncertainty grows. We observed that the difference between the robust and the corresponding nominal results becomes larger as the uncertainty grows. This observation demonstrates the advantage of the robust optimization formulation. In comparison with the original CVaR approach, our numerical examples indicates that the portfolio selection model based on the WCVaR performs robustly in practice.

The last numerical example tests the performance of the worst-case CVaR associated with mixture distribution uncertainty. We demonstrated that the worst-case requirement in the robust portfolio formulation does not influence the average performance of the portfolio considerably. Finally, we presented the comparison with an explicit solution of the robust portfolio model, which proposed the worst-case CVaR associated with mixture distribution uncertainty to be more suitable approach in the sense of return-risk trade-off.

Recently, a few researchers have paid more attention to general deviation measures that were introduced as an extension of standard deviation. An important deviation used in financial optimization, derived from CVaR, is CVaR deviation. Investigation of the worst-case CVaR deviation associated with different types of uncertainty and formulation of the corresponding portfolio optimization problems is left for further investigation.

Bibliography

[1] Artzner, P., F. Delbaen, J. M. Eber, D. Heath, 1999, Coherence measures of risk, Math. Finance 9: 203–228.

[2] Bawa, V. S., 1975, Optimal rules for ordering uncertain prospects, J. Financial Econom. 2: 95–121.

[3] Black, F., R. Litterman, 1992, Global portfolio optimization, Financial Analysts J. 48 28–43.

[4] Cantelli, F. P., 1910, Intorno ad un teorema fondamentale della teoria del rischio, Boll. Assoc. Attuar. Ital. (Milan): 1–23.

[5] Costa, O. L. V., A. C. Paiva, 2002, Robust portfolio selection using linear-matrix inequality, J. Econom. Dynam. Control 26 889–909.

[6] Chen Li, Simai He, Shuzhong Zhang, 2011, Tight Bounds for Some Risk Measures, with Applications to Robust Portfolio Selection, Operations Research 59: 847-865.

[7] Fan, K., 1953 Minimax theorems, Proc. National Acad. Sci. 39: 42–47.

[8] Fishburn, P. C., 1977, Mean-risk analysis with risk associated with below-target returns, Amer. Econom. Rev. 67: 116–126.

[9] Goldfarb, D., G. Iyengar, 2003, Robust portfolio selection problems, Operations Research 28 1–38.

[10] Lobo, M., L.Vandenberghe, S. Boyd, H. Lebret, 1998, Linear Algebra and its Applications, 284:193-228.

[11] Lobo, M. S., . Boyd, 2000, The worst-case risk of a portfolio. Technical report, Duke University.

[12] Lu, Z. 2006. A new cone programming approach for robust portfolio selection. Working paper, Department of Mathematics, Simon Fraser University, Burnaby, British Columbia, Canada.

[13] Markowitz, H. M., 1952, Portfolio Selection, The Journal of Finance 7(1), March.

[14] Ogryczak, W., A. Ruszczynski, 2002, Dual stochastic dominance and related mean-risk models, SIAM Journal on Optimization 13: 60–78.

[15] Rockafellar, R.T., S. Uryasev, 2000, Optimization of conditional value-at-risk, Journal of Risk 2, 21–41.

[16] Rockafellar, R. T., S. Uryasev, 2002, Conditional value-at-risk for general loss distributions, Journal of Banking and Finance 26: 1443–1471.

[17] Zhu, Shushang, Masao Fukushima, 2009, Worst-Case Conditional Value-at-Risk with Application to Robust Portfolio Management, Operations Research 57: 1155–1168.

List of Figures

4.1 Daily returns of AMZN 30
4.2 Optimal values of $\mathrm{CVaR}_{0.95}$ for different values of minimal required return r_0 31
4.3 Worst-case CVaRs of nominal optimal and robust optimal portfolios. 32
4.4 Daily returns of MGT 35

List of Tables

4.1 Estimated means, variances, and skewness of returns 30
4.2 Comparison of CVaRs and expected returns for different values of r_0 and α 31
4.3 Worst-case CVaR of nominal optimal and robust optimal portfolios for different values of $\boldsymbol{\eta}$ and r_0. 33
4.4 Out-of-sample total return of nominal optimal and robust optimal portfolios for different values of $\boldsymbol{\eta}$ and r_0. 33
4.5 Estimated means, variances, and skewness of returns in different periods. 34
4.6 Comparison of associated risk of nominal optimal and robust optimal portfolios. 36
4.7 Comparison of expected return of nominal optimal and robust optimal portfolios. 36
4.8 Mean and variance of nominal optimal and robust optimal portfolios (1600 sample points) 37
4.9 Weights of the assets in the optimal portfolio 37

Attachments

The attached CD contains the following folders with the corresponding files:

1. Data - market data used for the numerical applications (.txt format);
2. Source - source codes (R Files);
3. Thesis - the electronic version of the thesis (PDF format).