Optimal Portfolio Selection Under the Estimation Risk in Mean Return

by

Lei Zhu

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Master of Mathematics
in
Computer Science

Waterloo, Ontario, Canada, 2008

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

This thesis investigates robust techniques for mean-variance (MV) portfolio optimization problems under the estimation risk in mean return. We evaluate the performance of the optimal portfolios generated by the min-max robust MV portfolio optimization model. With an ellipsoidal uncertainty set based on the statistics of sample mean estimates, min-max robust portfolios equal to the ones from the standard MV model based on the nominal mean estimates but with larger risk aversion parameters. With an interval uncertainty set for mean return, min-max robust portfolios can vary significantly with the initial data used to generate the uncertainty set. In addition, by focusing on the worst-case scenario in the mean return uncertainty set, min-max robust portfolios can be too conservative and unable to achieve a high return. Adjusting the conservatism level of min-max robust portfolios can only be achieved by excluding poor mean return scenarios from the uncertainty set, which runs counter to the principle of min-max robustness. We propose a CVaR robust MV portfolio optimization model in which the estimation risk is measured by the Conditional Value-at-Risk (CVaR). We show that, using CVaR to quantify the estimation risk in mean return, the conservatism level of CVaR robust portfolios can be more naturally adjusted by gradually including better mean return scenarios. Moreover, we compare min-max robust portfolios (with an interval uncertainty set for mean return) and CVaR robust portfolios in terms of actual frontier variation, portfolio efficiency, and portfolio diversification. Finally, a computational method based on a smoothing technique is implemented to solve the optimization problem in the CVaR robust MV model. We numerically show that, compared with the quadratic programming (QP) approach, the smoothing approach is more computationally efficient for computing CVaR robust portfolios.

Acknowledgements

First and foremost, I would like to thank my supervisor, Yuying Li, for her guidance throughout the process of completing this work. It was Yuying who provided me the opportunity to study in the field of computational finance, and supported my interest in portfolio optimization and financial risk management. This thesis would not be possible without her patience, advice and encouragement.

I would also like to thank my readers, Prof. Peter Forsyth and Prof. Michael Best, for taking the time to review my thesis and providing suggestions and corrections. I would also like to express my gratitude to Prof. Tony Wirjanto for giving me valuable advice on my work.

Special thanks to my parents for their endless love, support and encouragement throughout all of my seven-year studies in Canada. I am forever indebted to my parents, Caixiao Liu and Qingfeng Zhu, who never stop giving me a world of love.

Last but not least, I would like to thank my friends in the Scientific Computation Lab for giving me their helps and support, and making my time more enjoyable at the university.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

The Markowitz mean-variance (MV) model has been used as the standard framework for optimal portfolio selection problems. However, due to the estimation risk in the MV model parameters (including the mean return and the covariance matrix of returns), the applicability of the MV model is limited. In particular, small differences in the estimates of mean return can result in large variations in the portfolio compositions; thus, the input parameters must be estimated very accurately. However, in reality accurate estimation of the mean return is notoriously difficult; estimation of the covariance matrix is relatively easier. For this reason we focus on, in this thesis, the estimation risk in mean return only, and investigate appropriate ways to take this estimation risk into account when using the MV model.

1.1 Problem Definition

In the min-max robust MV portfolio optimization model, MV model parameters are modeled as unknown, but belong to bounded uncertainty sets that contain all, or most, possible realizations of the uncertain parameters. To alleviate the sensitivity of the MV model to uncertain parameter estimates, min-max robust optimization yields the min-max robust portfolio that is optimal (MV efficient) with respect to the worst-case scenarios of the parameters in their uncertainty sets. Since an unknown parameter may have infinite number of possible scenarios, its uncertainty set typically corresponds to some confidence level $p \in [0, 1]$ with respect to an assumed distribution. In this regard, min-max robust optimization is a quantile-based approach, with the boundaries of an uncertainty set equal to certain quantile values for p.

One drawback with the min-max robust MV model is that, it entirely ignores the severity of the tail scenarios which occur with a probability of $1 - p$. Instead, it determines a min-max robust portfolio solely based on the single quantile value which corresponds to the worst sample scenario of a MV model parameter. Thus, the dependence on a single worst sample scenario makes a min-max robust portfolio quite sensitive to the initial data used to generate uncertainty sets. In particular, inappropriate boundaries of uncertainty sets can cause min-max robust optimization to be either too conservative or not conservative enough. In practice, it can be difficult to choose appropriate uncertainty sets.

Zhu et al. [33] have shown that, with an ellipsoidal uncertainty set based on the statistics of sample mean estimates, the robust portfolio from the min-max robust MV model equals to the optimal portfolio from the standard MV model based on the nominal mean estimate but with a larger risk aversion parameter. Therefore, we focus on illustrating the characteristics of min-max robust portfolios with an interval uncertainty set. If the uncertainty interval for mean return contains the worst sample scenario, the min-max robust MV model often produces portfolios with very low returns. Portfolios with higher returns can be generated in the model by choosing the uncertainty interval to correspond to a smaller confidence interval. Unfortunately, this is at the expense of ignoring worse sample scenarios and runs counter to the principle of min-max robustness.

1.2 Thesis Contributions

In this thesis, we focus on the uncertainty of mean return, and propose a CVaR robust MV portfolio optimization model which determines a CVaR robust portfolio that is optimal (MV efficient) under the estimation risk in mean return. The CVaR robust MV model uses the Conditional Value-at-Risk (CVaR) to measure the estimation risk in mean return, and control the conservatism level of a CVaR robust portfolio with respect to estimation risk by adjusting the confidence level of CVaR, $\beta \in [0, 1)$. As a risk measure, CVaR is coherent, see Artzner et al. [2], and can be used to quantify the risk of a portfolio under a given distribution assumption. In the traditional return-risk analysis, CVaR is used to quantify the portfolio loss due to the volatility of asset returns. In the estimation risk analysis addressed in this thesis, CVaR is used to quantify the portfolio mean loss (which is a function of portfolio expected return) due to mean return uncertainty. In this regard, the CVaR of a portfolio's mean loss is used as a performance measure of this portfolio under the estimation risk in mean return.

Instead of focusing on the worst sample scenario in the uncertainty set of mean return, the CVaR robust MV model determines an optimal portfolio based on the tail of the portfolio's mean loss scenarios (with respect to an assumed distribution) specified by the confidence level β. In addition, the conservatism level of the portfolio with respect to the estimation risk in mean return can be adjusted by changing the value of β. As β approaches 1, the CVaR robust MV model considers the worst mean loss scenario and the resulting portfolio is the most conservative. As the value of β decreases, better mean loss scenarios are included for consideration and the dependency on the worst case is decreased. Thus the resulting portfolio is less conservative. When $\beta = 0$, all sample mean loss scenarios are considered in the model; this may be appropriate when an investor has complete tolerance to estimation risk. Thus the confidence level β can be interpreted as an estimation risk aversion parameter.

Diversification reduces the overall portfolio return risk by spreading the total investment across a wide variety of asset classes. We illustrate that, no matter how an interval uncertainty set is selected to achieve the desired level of conservatism, the maximum worst-case expected return portfolio from the min-max robust MV model (i.e., the risk aversion parameter $\lambda = 0$) typically consists of a single asset. In contrast, the maximum CVaR expected return portfolio can consist of multiple assets. In addition, we computationally show that the diversification level of CVaR robust portfolios decreases as the value of β (which is interpreted as an estimation risk aversion parameter) decreases. We also consider two different distributions to characterize the uncertainty in mean return, and compare the diversification level of CVaR robust portfolios between two different sampling techniques.

One way of computing CVaR robust portfolios is to discretize, via simulation, the CVaR robust optimization problem. This can be formulated as a quadratic programming (QP) problem, where the CVaR function is approximated by a piecewise linear function. However, the QP approach becomes inefficient when the scale of the optimization problem becomes large. As an alternative of the QP approach, a computational method based on a smoothing technique is implemented to compute CVaR robust portfolios. Differently from the QP approach, the smoothing approach uses a continuously differentiable piecewise quadratic function to approximate the CVaR function. Comparisons on computational efficiency and approximation accuracy are made between the two approaches when they are applied in the CVaR robust MV model. We show that the smoothing approach is more computationally efficient, and can provide sufficiently accurate solutions when the number of scenarios becomes large.

1.3 Thesis Organization

This thesis is organized as follows: Chapter 2 introduces the standard MV model and demonstrates the estimation risk in mean return for the model. This chapter also discusses the various techniques proposed in current literatures to combat the impact of estimation error.

Chapter 3 reviews the min-max robust MV portfolio optimization model and highlights its potential problems. We discuss the sensitivity of min-max robust portfolios to the initial return samples which generate the uncertain intervals. In addition, we consider a variance-based technique to produce portfolios which are less sensitive to the initial data, and emphasize the importance of being able to achieve a high expected return in a robust approach.

Chapter 4 presents the CVaR robust MV portfolio optimization model. We show how this model adjusts a portfolio's conservatism level with respect to the estimation risk in mean return.

Chapter 5 computationally compares the characteristics of the actual frontiers generated by the min-max robust (for an interval uncertainty set of mean return) and the CVaR robust MV models in terms of actual frontier variation, portfolio efficiency, and portfolio diversification.

Chapter 6 addresses the computational efficiency issue for computing CVaR robust portfolios. We show that a smoothing approach proposed in [1] is significantly more efficient than the QP approach for computing CVaR robust portfolios. In addition, the solution obtained by the smoothing approach can be very close to that obtained by the QP approach when the number of scenarios becomes large.

Chapter 7 concludes the thesis by presenting the research achievements and indicating the areas that could benefit from further study.

Chapter 2

Mean-Variance Portfolio Optimization and Estimation Risk

This chapter provides the background knowledge for this thesis. It starts with the formal definition of the Markowitz mean-variance (MV) model. Then it illustrates the estimation risk of the MV model. Finally, it discusses the various techniques proposed in recent research to combat the impact of estimation error.

2.1 Markowitz Mean-Variance Model

Portfolio optimization is used in financial portfolio selection to maximize return and minimize risk. In the mean-variance (MV) portfolio optimization model introduced by Markowitz [21], the portfolio return is measured by the expected rate of the random portfolio return, and the associated risk is measured by the variance of the return.

2.1.1 Mathematical Notations

Assume that a rational investor makes investment decisions for a portfolio that contains n assets. Let $\mu \in R^n$ be the mean vector with μ_i as the mean return of asset i, $1 \leq i \leq n$, and $x \in R^n$ be the decision vector with x_i as the proportion of holding in the i^{th} asset . The portfolio expected return μ_p is the weighted average of individual asset return and can be defined as:

$$\mu_p = \sum_{i=1}^{n} x_i \mu_i \ . \tag{2.1}$$

The variance and covariance of individual assets are characterized by a n-by-n positive semi-definite matrix Q, such that:

$$Q = \begin{bmatrix} \sigma_{11} & \dots & \sigma_{1n} \\ \vdots & \ddots & \vdots \\ \sigma_{n1} & \dots & \sigma_{nn} \end{bmatrix} , \tag{2.2}$$

where σ_{ii} is the variance of asset i, and σ_{ij} is the covariance between asset i and asset j. Therefore, the variance of portfolio return, σ_p^2, can be calculated by:

$$\sigma_p^2 = x^T Q x = \sum_{i=1}^{n} \sum_{j=1}^{n} x_i x_j \sigma_{ij}. \tag{2.3}$$

2.1.2 Model Definition

The MV model assumes that, for a given level of risk (measured by variance), a rational investor would choose the portfolio with the highest expected return; similarly, for a given level of expected return, a rational investor would choose the portfolio with the lowest risk. In other words, a portfolio is said to be optimal (MV efficient) if there is no portfolio having the same risk with a greater expected return, and there is no portfolio having the same expected return with a lower risk. Therefore, the MV model can be formulated mathematically as three equivalent optimization problems:

(1) Maximizing the expected return for a upper limit on variance:

$$\begin{aligned} \max_{x} \quad & \mu^T x \\ \text{s.t.} \quad & x^T Q x \le V \\ & x \in \Omega \end{aligned} \tag{2.4}$$

(2) Minimizing the variance for a lower limit on expected return:

$$\begin{aligned} \min_{x} \quad & x^T Q x \\ \text{s.t.} \quad & \mu^T x \ge R \\ & x \in \Omega \end{aligned} \tag{2.5}$$

(3) Maximizing the risk-adjusted expected return:

$$\begin{aligned} \min_{x} \quad & -\mu^T x + \lambda x^T Q x \\ \text{s.t.} \quad & x \in \Omega, \end{aligned} \tag{2.6}$$

where $\lambda \ge 0$ is the risk-aversion parameter which measures how the investor views the trade-off between risk (which is measured variance) and expected return. The symbol Ω used in the above three problems denotes the additional linear constraints for the feasible portfolio sets, e.g.,

$$\Omega = \{x \in \mathbb{R}^n \mid \sum_{i=1}^{n} x_i = 1, x \ge 0\}, \tag{2.7}$$

which corresponds to the case where no short-sales are allowed, and all available money for investment is allocated to the n assets. Note that here x_i denotes the proportion of holding of the i^{th} asset.

Formulation Equivalence

Problem (2.4) maximizes a (concave) linear function subject to quadratic and linear constraints; while problem (2.5) and (2.6) minimize convex quadratic functions subject to linear constraints. When μ is not a multiple of a vector that contains n ones, the three problems can be mathematically equivalent, i.e., an optimal solution $x^*(\lambda)$ of problem (2.6) is also an optimal solution of problem (2.5) such that $\mu^T x^*(\lambda) = R$ for some R, and similarly, is an optimal solution of problem (2.4) such that $x^*(\lambda)^T Q x^*(\lambda) = V$ for some V. Problem (2.5) and (2.6) are commonly used in practice as they are both formulated as convex quadratic programming (QP) problems and can be efficiently solved using readily available optimization software.

Risk-Aversion Parameter

The risk-aversion parameter λ used in problem (2.6) represents the degree with which investors want to maximize return at the expense of assuming more risk. Each investor is willing to take a certain amount of risk to get a level of expected return. Since return is compensated by risk, investors have to balance the trade-off between return and risk by using appropriate λ values. As the value of λ decreases, investors focus more on maximizing expected return than minimizing risk. In this case, both the expected return and the associated risk will increase. There are also two extreme situations where investors only care about maximizing return and minimizing risk: when $\lambda = 0$, problem (2.6) gives us the maximum-return portfolio without considering the associated risk. On the other hand, when $\lambda = \infty$, problem (2.6) gives us the minimum-variance portfolio without considering the expected return.

Linear Constraints

In the MV model (2.6), only the budget constraint and the no-shortselling constraint are specified in Ω. However, in real investment practice, there may be other linear constraints that need to be considered such as transaction costs and trading size limits on certain assets. Therefore, we can extend the MV model (2.6) to the following generalization:

$$\begin{aligned} \min_{x} \quad & -\mu^T x + \lambda x^T Q x \\ \text{s.t.} \quad & Cx \leq d, \\ & Ex = v, \\ & x \geq 0, \end{aligned} \tag{2.8}$$

where $C \in R^{m \times n}$, $E \in R^{m \times n}$, $d \in R^m$ and $v \in R^m$. The inequality $Cx \leq d$ and the equality $Ex = v$ can be used to express the linear constraints mentioned above.

2.1.3 Efficient Frontier

By solving problem (2.6) for all possible values of λ from 0 to ∞, we can obtain the efficient frontier: it contains the entire set of MV efficient portfolios ranging from the maximum expected return to the minimum variance. The same efficient frontier can also be generated by solving problem (2.4) for all possible values of V, or by solving problem (2.5) for all possible values of R. Any point in the region below the efficient frontier is not MV efficient, since there is another portfolio with the same risk and a higher expected return, or with the same expected return and a lower risk.

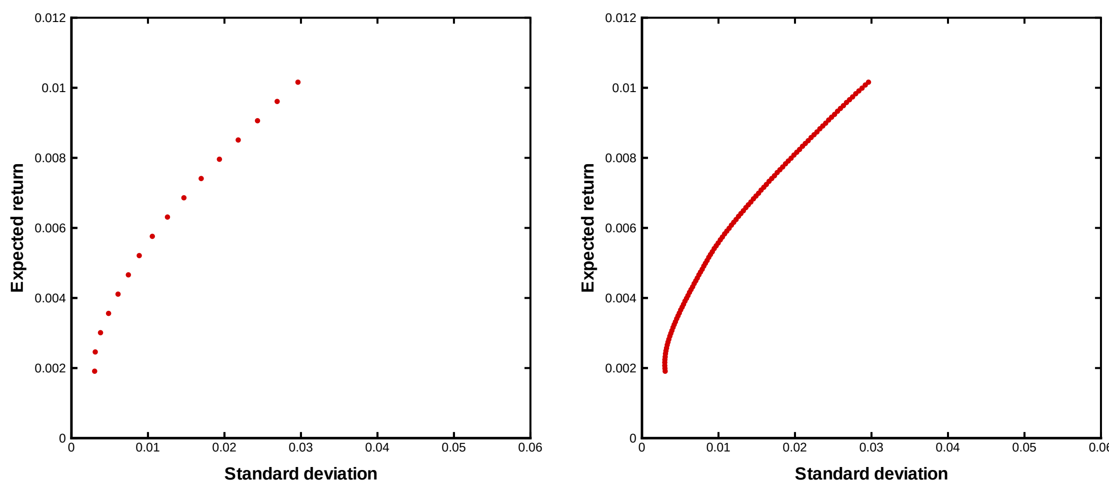
Since it is impossible to generate infinite number of portfolios, we must approximate the exact efficient frontier with a finite algorithm. Here we consider an algorithm which generates the efficient frontier by first computing the portfolios with the maximum and the minimum expected returns, and then solving problem (2.5) subject to a finite number of expected returns that lie between the two extreme points. Let x_{min} and x_{max} denote the portfolios that achieve the minimum expected return, R_{min}, and the maximum expected return, R_{max}, respectively. This algorithm can be described as the following:

Algorithm 1 Generating Efficient Frontier

1. Compute x_{max} by solving problem (2.6) with $\lambda = 0$. Then set $R_{\text{max}} = \mu^T x_{\text{max}}$.
2. Compute x_{min} by solving problem (2.5) without the expected return constraint. Then set $R_{\text{min}} = \mu^T x_{\text{min}}$.
3. Generate m equally spaced values between R_{min} and R_{max} such that: $R_{\text{min}} \leq R_1 \leq R_2 \ldots \leq R_{\text{m}} \leq R_{\text{max}}$. For each $i \in m$, compute x_{i} by solving problem (2.5) with R_{i} as the expected return constraint.

Step 1 generates the optimal portfolio that has the maximum expected return without considering the associated risk; while Step 2 generates the optimal portfolio that has the minimum risk without considering the expected return; this resulting portfolio also has the minimum expected return otherwise it would not be MV efficient. Having determined the maximum and the minimum expected returns in the previous two steps, Step 3 generates m optimal portfolios whose expected returns equally lie in between the two extreme values. The larger the value of m, the better approximation obtained for the efficient frontier. Once obtain the portfolio weights x_{min}, x_1, x_2, …, x_{m}, x_{max} from Algorithm 1, we can approximate the exact efficient frontier by plotting these points in a two-dimensional space with the standard deviation (horizontal axis) and the expected return (vertical axis). Figure 2.1(a) and Figure 2.1(b) depict the approximated efficient frontier generated by using Algorithm 1 with $m = 16$ and $m = 100$ respectively. As we can expect, when $m \to \infty$, the resulting efficient frontier will approximate the exact one.

We just illustrate in Algorithm 1 a simple algorithm to approximate the exact efficient frontier. Using equally spaced points for approximation may possibly miss some important “intervals” on the efficient frontier, and solving an individual QP problem for each of the point can be computationally expensive when the number of points becomes large. Alternative algorithms can be applied to generate the efficient frontier more accurately and efficiently. For example, Markowitz [22] introduces a 'critical line' algorithm in a form of parametric quadratic programing (QP). This algorithm iteratively traces out the efficient frontier by identifying the 'corner' portfolios, which are the points where a stock either enters or leaves the current portfolios. Therefore, by only determining the 'corner' portfolios, the computational cost for generating the efficient frontier can be dramatically decreased. Note that the 'critical line' algorithm requires the covariance matrix Q to be positive definite. However, Best [4] proposes an algorithm for solving the parametric QP problem such that Q is required to be positive semi-definite only.



(a) 16 discretization points

(b) 100 discretization points

Figure 2.1: Approximated efficient frontiers generated using Algorithm 1 for the 8-asset example in Table 2.1.

2.2 Estimation Risk in MV Model Parameters

Despite its theoretical importance in modern finance, the MV model is known to have severe performance limitations in practice. One of the basic problems that limits the applicability of the MV model is the estimation error of the input parameters, i.e., asset mean returns and the covariance matrix of returns. Michaud [25] discusses the implications of estimation error for portfolio managers. Best and Grauer [5] analyze the effect of changes in the mean return of assets on the MV efficient frontier and the composition of optimal portfolios. Chopra and Ziemba [9] analyze the impact of errors in means, variances and covariances on investor's utility function, and study the relative importance of these errors. Broadie [7] investigates the effect of errors in parameter estimates on the results of actual frontiers, which are obtained by applying the true parameters on the portfolio weights derived from their estimated values. All of these studies show that different input estimates to the MV model result in large variations in the composition of MV efficient portfolios. Since, in reality, accurate estimation of input parameters is a very difficult task, the estimation risk introduced by estimation error must be taken into account when using the MV model.

2.2.1 MV Model Under Estimation Risk

The parameters of the MV model are the asset mean returns and the covariance matrix of returns, which are denoted as μ and Q respectively. To implement the MV model in practice, one may estimate these parameters based on empirical return samples. Let the estimated mean returns and covariance matrix be $\bar{\mu}$ and $\bar{Q}$ respectively. Using the estimated parameters, the actual portfolio optimization problem becomes

$$\begin{aligned} \min_{x} \quad & -\bar{\mu}^T x + \lambda x^T \bar{Q} x \\ \text{s.t.} \quad & x \in \Omega. \end{aligned} \tag{2.9}$$

The solution of problem (2.9) coincides with the one of problem (2.6) only if $\mu = \bar{\mu}$ and $Q = \bar{Q}$. However, due to the estimation error introduced in the estimation process, the estimated parameters (especially $\bar{\mu}$) can have large errors. Therefore, the resulting portfolio weights computed from problem (2.9) fluctuate substantially for different μ estimates, and the out-of-sample performance of these portfolios can be quite poor.

2.2.2 An Example of Estimation Risk

To demonstrate the effect of the estimation error on the computation of MV efficient frontiers, we conduct the following experiment. Suppose there are eight risky assets and their true parameters, the means μ and the covariance matrix Q, are given in Table 2.1. Assume that the asset returns constitute a joint normal distribution, we generate, from μ, 48 return samples using Monte Carlo simulation (we can consider the samples as 48 monthly returns of the eight assets). From these samples, we calculate the sample means $\bar{\mu}$ and the covariance matrix $\bar{Q}$. These two estimated parameters are given in Table 2.2.

Comparing the values between Table 2.1 and Table 2.2, we find that the estimation error of $\bar{Q}$ is relatively small. The entry with the largest absolute estimation error in $\bar{Q}$ is $\bar{Q}_{66}$, for which the value is 0.0931×10^{-2}. With Q_{66} equals to 0.2691×10^{-2}, the relative estimation error, which is the ratio (absolute value) between the absolute error and the true value, is 0.34. On the other hand, the estimated mean returns in $\bar{\mu}$ have much larger errors. The assets with the highest and the lowest mean return values in μ are Asset1 and Asset7 respectively, for which μ_1=1.016×10^{-2} and μ_7=-0.112×10^{-2}; while the corresponding assets in $\bar{\mu}$ are Asset3 and Asset6, for which $\bar{\mu}_3$=1.8032×10^{-2} and $\bar{\mu}_6$=-0.4775×10^{-2}. The entry with the largest absolute estimation error in $\bar{\mu}$ is $\bar{\mu}_3$, for which the value is 1.3472×10^{-2}. In addition, the relative absolute estimation error of $\bar{\mu}_3$ is 2.79, which is about three times of the true value μ_3.

Table 2.1 True mean vector and covariance matrix

Mean Return Vector μ

$10^{-2}\times$	Asset1	Asset2	Asset3	Asset4	Asset5	Asset6	Asset7	Asset8
	1.0160	0.47460	0.47560	0.47340	0.47420	-0.0500	-0.1120	0.0360

Covariance Matrix Q

$10^{-2}\times$	Asset1	Asset2	Asset3	Asset4	Asset5	Asset6	Asset7	Asset8
Asset1	0.0980							
Asset2	0.0659	0.1549						
Asset3	0.0714	0.0911	0.2738					
Asset4	0.0105	0.0058	-0.0062	0.0097				
Asset5	0.0058	0.0379	-0.0116	0.0082	0.0461			
Asset6	-0.0236	-0.0260	0.0083	-0.0215	-0.0315	0.2691		
Asset7	-0.0164	0.0079	0.0059	-0.0003	0.0076	-0.0080	0.0925	
Asset8	0.0004	-0.0248	0.0077	-0.0026	-0.0304	0.0159	-0.0095	0.0245

Table 2.2 Estimated mean vector and covariance matrix for the data in Table 2.1

Estimated Mean Return Vector $\bar{\mu}$

$10^{-2}\times$	Asset1	Asset2	Asset3	Asset4	Asset5	Asset6	Asset7	Asset8
	1.6517	1.5015	1.8032	0.5551	0.8783	-0.4775	0.1350	-0.1492

Estimated Covariance Matrix $\bar{Q}$

$10^{-2}\times$	Asset1	Asset2	Asset3	Asset4	Asset5	Asset6	Asset7	Asset8
Asset1	0.0707							
Asset2	0.0394	0.1185						
Asset3	0.0312	0.0467	0.2432					
Asset4	0.0064	0.0049	-0.0196	0.0097				
Asset5	-0.0023	0.0256	-0.0141	0.0038	0.0319			
Asset6	-0.0130	-0.0095	0.0121	-0.0089	-0.0158	0.1760		
Asset7	-0.0093	0.0147	0.0245	-0.0062	0.0115	-0.0500	0.1015	
Asset8	0.0089	-0.0144	0.0095	0.0008	-0.0231	0.0150	-0.0158	0.0215

2.2.3 Visualizing Estimation Risk

The effect of the estimation error on the computation of efficient frontier can be observed from Figure 2.2. With the estimated parameters $\bar{\mu}$ and $\bar{Q}$ from Table 2.2, we can compute a sequence of optimal portfolio weights using Algorithm 1; by plotting these weights with the true parameters μ and Q, we obtain a frontier. This frontier reflects how the portfolios obtained from the estimated parameters really behave based on the true parameters, and is defined by Broadie [7] as the actual frontier.

Observed from Figure 2.2, the actual frontier is clearly below the true efficient frontier. As the risk-aversion parameter λ decreases, the investment is focused more on maximizing expected return than minimizing risk. This leads to less diversified portfolios for which the estimation error can be more significant, especially when the estimated highest-return asset is different from the true one. For example, prior to and include point A, all portfolios on the actual frontier consist of at least three different assets; this diversification reduces the impact of estimation error on μ. However, the portfolios between point A and point B consist of only Asset1 and Asset3, and since Asset3 has the highest return in $\bar{\mu}$, its proportion of holding is gradually increased from A to B. In particular, when setting $\lambda = 0$ in problem (2.9), all investment is allocated to Asset3 for achieving the maximum portfolio return. However, the asset with the true highest return in μ is Asset1 and μ_1 is much higher than μ_3, reducing the proportion of holding of Asset1 but increasing the one for Asset3 causes the actual frontier from point A to point B appear downward. We also observe that the distance between the actual frontier and the true efficient frontier decreases as the portfolio risk decreases. i.e., the maximum expected return portfolios are relatively far away from each other while the minimum risk portfolios are quite close. This coincides with the experimental results in Broadie [7]: covariance matrix is much easier to estimate than mean returns.

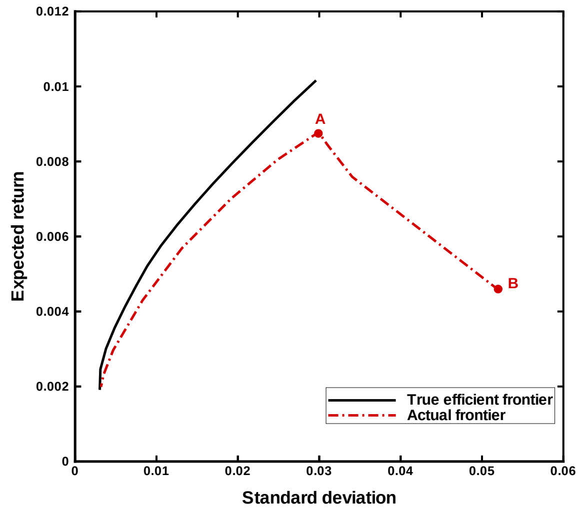
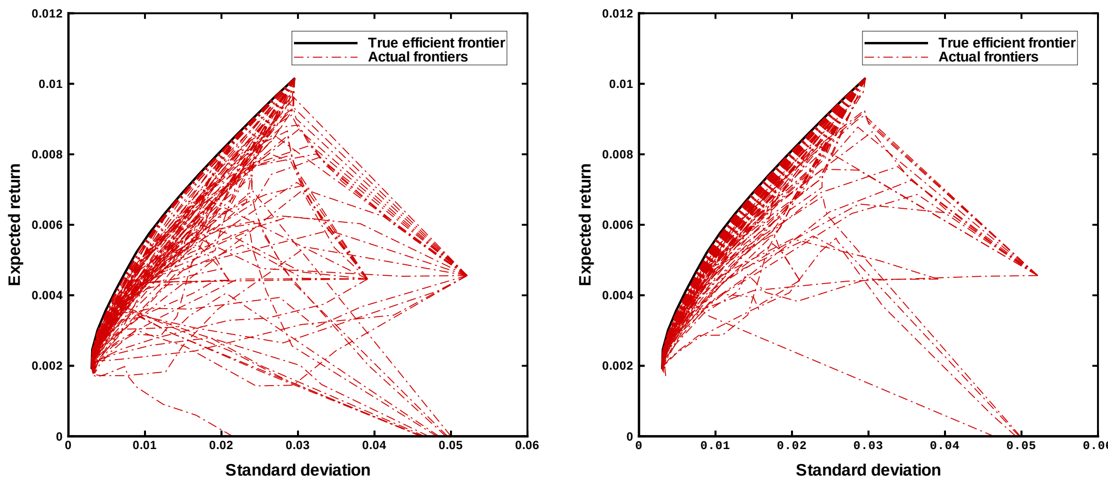


Figure 2.2: True efficient frontier and actual frontier using 48 simulated monthly returns.

2.2.4 Estimation Risk vs. Stationarity

The values in Table 2.2 are estimated based on 48 simulated monthly returns. The following example shows that estimation error can decrease as the number of simulated returns increases. We repeat the above estimation process 100 times using 48 months of simulated data, and plot the actual frontiers obtained during each process in Figure 2.3(a). Next, we re-produce the actual frontiers using 96 months of simulated data and plot them in Figure 2.3(b). The difference between the two plots depicts that the performance of actual frontiers are improved with more data, i.e., comparing to Figure 2.3(a), the actual frontiers in Figure 2.3(b) become closer to the true efficient frontier, and their variation is much smaller. One may suggest increasing the accuracy of estimated parameters by using more data. However, this is difficult to be achieved in practice. First, large amount of historical data might not be available to be used for estimation. Second, using very old historical data makes it difficult to assume stationarity on the estimated parameters; see Broadie [7]. Therefore, there is a trade-off between maintaining stationarity and reducing estimation error when deciding the amount of data used for estimation.



(a) 48 months of simulated data

(b) 96 months of simulated data

Figure 2.3: True efficient frontier and actual frontiers for the 8-asset example in Table 2.1.

2.3 Related Work

2.3.1 Robust Optimization

Various techniques have been proposed to reduce the impact of estimation error, and robust portfolio optimization is an active research area; see e.g., Goldfarb and Iyengar [14], Tütüncü and Koenig [30], and Garlappi et al. [12]. In the robust optimization framework introduced by these papers, input parameters are modeled as unknown, but belong to bounded uncertainty sets that contain all, or most, values of the uncertain inputs. Therefore, robust optimization determines the optimal portfolio under the worst-case scenario of the inputs in their uncertainty sets. Robust optimization provides a conservative framework to determine an optimal portfolio under model parameter uncertainty. However, such a framework tends to be too pessimistic and unable to achieve high portfolio returns, especially for less risk-averse investors. In addition, the solution provided by this framework can be very sensitive to the choice of uncertainty sets. Chapter 3 addresses these issues, and presents more detail discussion on the robust optimization approach.

2.3.2 Robust Estimation

Another related approach is robust portfolio estimation. Unlike robust optimization, which defines the unknown parameters as uncertainty sets and determines optimal portfolios under the worst-case performance, robust estimation is based on a single point estimate which is generated by a robust estimator. A standard framework adopted in this approach is the Bayesian estimation. In the Bayesian framework, an investor is assumed to have a pre-specified prior, which is the subjective view on the distribution of returns. The predictive distribution of returns is therefore calculated based on the prior and the confidence that the investor has on this prior. The more confidence on the prior, the more subjective the predictive distribution is. Many applications of the Bayesian method have been used for robust portfolio selection problems. Jorion [18] uses a empirical Bayesian framework to develop a shrinkage estimator of the mean returns under estimation and model risk. Black and Litterman [6] propose a Bayesian approach that combines the investor's subjective views and the implied returns which are determined based on market equilibrium. Ľ. Pástor and Stambaugh [32] form the prior by incorporating investors' degree of belief in the Capital Asset Pricing Model (CAPM). In the above Bayesian models, the optimal portfolio is determined by maximizing the expected utility of an investor, where the expectation is taken with respect to the predictive distribution.

2.3.3 Robust Statistics

Another important technique to generate a robust estimator is robust statistics. In classical statistics, estimation methods rely heavily on the assumptions that may not be met in practice. As a result, when data outliers exist or data distribution assumptions are violated, the performance of these methods is often quite poor. On the other hand, as an extension of the classical statistics, robust statistics take into account the possibility of outliers or deviation from distribution assumptions. In particular, robust statistics can be used to generate the robust estimators that provide meaningful information even when empirical statistical assumptions are different from the assumed ones. Many applications which are based on robust statistics have been proposed for robust portfolio selection. However, the robust estimators used by these applications are quite different. For examples, Cavadini et al. [8] use the minimum covariance determinant (MCD) estimator, Vaz-de Melo and Camara [31] use M-estimators, and Perret-Gentil and Victoria-Feser [28] use S-estimators. The determination of a robust portfolio based on robust statistics usually takes two steps. First, estimates of the unknown parameters are determined by robust estimators. Second, robust portfolios are computed by solving the MV optimization problem which takes the robust estimates as inputs. However, DeMiguel and Nogales [11] propose an approach where both data estimation and portfolio optimization are preformed in one step. In that paper, a robust portfolio is determined by minimizing the risk estimated by M-estimators, and the risk minimization problem is solved via a single nonlinear problem.

2.3.4 Other Approaches

Besides robust optimization and robust estimation, a number of other approaches have been proposed to address the sensitivity of MV portfolios to model parameter uncertainty. One popular approach to increase the stability of a portfolio is to place constraints on the amount of an asset can have in the portfolio. Chopra and Ziemba [9] suggest that the solution of the portfolio optimization, which is subject to portfolio weight constraints, has better performance than the one without the constraints. Jagannathan and Ma [16] propose imposing short-selling constraints, and show that this can reduce the impact of estimation error on the stability and the performance of the minimum-variance portfolios. Instead of running a single robust portfolio optimization, Michaud [26] proposes the Resampled Efficiency (RE) optimization technique, which finds an optimal portfolio by averaging the portfolio weights obtained from different simulations. By conducting simulation performance test, they show that the RE optimized portfolios not only outperform the classical MV optimized portfolio but also give a smoother transition as portfolio return requirements change. Garlappi et al. [12] extend the MV model to a multi-prior model where mean returns are obtained by using maximum likelihood estimation. Unlike the Bayesian approaches which use a single prior and assume the investor is neutral to uncertainty, the multi-prior model allows for multiple priors and aversion to uncertainty. Their analysis suggests that, for both the international and the domestic data considered, the portfolios generated by the multi-prior model have better out-of-sample performance (such as in Sharpe ratio and portfolio mean-standard deviation ratio) than that generated by the Bayesian approaches. In addition, compared with the MV model which does not take parameter uncertainty into account, the multi-prior model reduces the fluctuation of portfolio weights over time.

All the approaches mentioned above compute robust portfolios under the estimation risk of the MV model parameters; however they preserve robustness from different perspectives. For example, robust estimation focuses on improving the estimation of optimization inputs; while the RE technique focuses on enhancing the optimizer. Note that these approaches are not mutually exclusive from each other, but could be used in conjunctions. For examples, one could use a robust estimator generated using robust statistics to determine uncertainty sets, from which the worst-case parameters are chosen for robust optimization; Michaud [26] suggests using Bayesian approach to improve the estimation of the inputs used by the RE optimizer.

2.4 Conclusion and Remarks

This chapter gives the background knowledge that are relevant to the ideas discussed in the remaining parts of the thesis. In Section 2.3, we briefly review the approaches proposed to deal with the estimation risk of the MV model. In next chapter, we focus on the robust optimization approach, which is one of the most important approaches in this research area.

Chapter 3

Min-max Robust Mean-Variance Portfolio Optimization

This chapter reviews the min-max robust portfolio optimization framework and highlights its potential weakness. We focus on the min-max robust MV model with interval uncertainty sets and analyze the performance of the resulting min-max robust portfolios. We also discuss some general criterion that should be used for evaluating the performance of robust portfolios.

3.1 Robust Portfolio Optimization

Robust optimization is an approach for solving optimization problems in which the data is uncertain and is only known to belong to some bounded uncertainty set. Consider the general optimization problem:

$$\begin{aligned} \min_x \quad & f(x,\xi) \\ \text{s.t.} \quad & F(x,\xi) \leq 0, \end{aligned} \tag{3.1}$$

where ξ is the data element of the problem, x $\in R^n$ is the decision vector, and $F(x,\xi) \in R^m$ are m constraint functions. For deterministic optimization problems, ξ is assumed to be known and fixed. However, in reality ξ may be uncertain but belong to a given uncertain set U. In this case, the optimal solution x must both satisfy the constraints for every possible realization of ξ in U, and give the best possible guaranteed value of the objective under the worst-case of ξ. Therefore, the robust counterpart of the optimization problem (3.1) can be formulated as:

$$\begin{aligned} \min_x \quad & \sup_{\xi \in U} f(x,\xi) \\ \text{s.t.} \quad & F(x,\xi) \leq 0, \quad \forall \xi \in U \ . \end{aligned} \tag{3.2}$$

The uncertainty set U contains all, or most, possible scenarios of ξ, and can be represented by various structures such as intervals (box) or ellipsoids. Depending on the structures of the uncertainty set being used, we can obtain different robust counterparts for the same optimization problem. For example, Ben-Tal and Nemirovski [3] show that when U is an ellipsoid uncertainty set, the robust counterpart for a LP problem is a conic quadratic problem, and the one for a QP problem is a semi-definite program. Both of these robust counterparts are tractable problems that can be solved using efficient algorithms such as interior-point methods.

3.1.1 Min-max Robust MV Model

An important application of robust optimization is to compute optimal (MV efficient) portfolios under the uncertainty of MV model parameters. As mentioned in Section 2.2, the parameters (including the means μ and the covariance matrix Q) for the MV model (2.6) are unknown, and using the estimates of these parameters leads to an estimation risk in portfolio selection. In particular, small differences in the estimate of μ can result in large variations in the composition of an optimal portfolio. To alleviate the sensitivity of the MV model to the parameter estimates, robust optimization is applied to determine optimal portfolios under the worst-case scenario of the parameters in their uncertainty sets. These uncertainty sets often correspond to certain confidence levels under an assumed distribution. Mathematically, the corresponding robust formulation for the MV model (2.6) can be expressed as the following problem:

$$\begin{array}{ll} \min\limits_{x} & \max\limits_{\mu \in S_\mu, Q \in S_Q} -\mu^T x + \lambda x^T Q x \\ \text{s.t.} & x \in \Omega, \end{array} \tag{3.3}$$

where S_μ and S_Q are the uncertainty sets for μ and Q, respectively. Problem (3.3) is often referred as the min-max robust MV portfolio optimization model, as the optimal solution x minimizes the maximum case (given by the selected μ and Q) of the objective function.

A closely related robust version of the MV model is proposed by Goldfarb and Iyengar [14]. Unlike the problem (3.3), which minimizes the worse-case risk-adjusted expected return, it minimizes the worst-case variance of the portfolio subject to the worst-case expected return constraint. i.e., to solve the following problem:

$$\begin{aligned} \min_{x} \quad & \max_{Q \in S_Q} x^T Q x \\ \text{s.t.} \quad & \min_{\mu \in S_\mu} \mu^T x \geq R, \\ & x \in \Omega, \end{aligned} \tag{3.4}$$

where R is the pre-specified lower bound of the worst-case expected return. Tütüncü and Koenig [30] show that an optimal solution $x^*(\lambda)$ of the problem (3.3) is also an optimal solution of problem (3.4) when $R = \min_{\mu \in S_\mu} \mu^T x^*(\lambda)$ for some λ and R.

The uncertainty sets S_μ and S_Q in the above robust portfolio optimization problems can be represented in different ways. Goldfarb and Iyengar [14] use ellipsoidal constraints to describe uncertainty sets, and formulate problem (3.3) as a second-order cone programming (SOCP) problem. Tütüncü and Koenig [30] consider uncertainty sets as intervals, and solve problem (3.3) using a saddle-point method. In addition, Lobo and Boyd [20] show that an optimal portfolio that minimizes the worst-case risk under each or a combination of the above uncertainty structures can be computed efficiently using analytic center cutting plane methods.

3.1.2 Other Robust Models

In addition to problem (3.3) and problem (3.4), various other robust portfolio optimization problems have been proposed by recent research. For examples, the dual of (3.4) is the robust maximum return problem, which maximizes the worst-case expected return subject to a constraint on the worst-case variance; as an alternative to minimizing the worst-case variance as in problem (3.4), one can choose a portfolio that maximizes the worst-case ratio of the expected excess return on the portfolio, i.e. the ratio of the return in excess of the risk-free rate to the standard deviation of the return; see Goldfarb and Iyengar [14]. Recently, VaR and CVaR risk measures have often been used to replace variance in robust portfolio optimization applications. Ghaoui et al. [13] propose a robust approach that minimizes the worst-case VaR when the return distribution is partially unknown, and cast the optimization problems as semi-definite programs. Zhu and Fukushima [34] consider the minimization of the worst-case CVaR under both box uncertainty and ellipsoidal uncertainty, and cast the corresponding problems as LP programs and SOCP programs, respectively.

3.2 Min-max Robust MV Portfolio Optimization

In this section, we analyze the characteristics of the min-max robust actual frontier, which is formed by the robust portfolios computed using the min-max robust MV model (3.3). Similar to [7], we consider the min-max robust actual frontier in the space of mean and variance: the portfolio expected return and variance are computed by applying the true parameter values μ and Q on the min-max robust portfolios obtained from solving problem (3.3) for different values of $\lambda \geq 0$. It represents how the min-max robust portfolios, which are determined under the worst sample scenarios of MV model parameters, actually behave when applied with the true parameter values.

We begin our discussion of min-max robust optimization using an ellipsoidal structure for the uncertainty set of mean return. We show that the resulting min-max robust portfolios actually equal to those computed from the standard MV model based on the nominal mean estimates but with a larger risk aversion parameter. Subsequently, we focus on the min-max robust optimization which uses an interval structure for the uncertainty set of mean return.

3.2.1 Ellipsoidal Uncertainty Set

We consider the ellipsoidal uncertainty set that is based on the following statistical properties of mean estimates. Assume that asset returns have a joint normal distribution, and mean estimate $\bar{\mu}$, which is estimated jointly for all assets, is computed from T samples of n assets. If the covariance matrix Q is known, then the quantity

$$\frac{T(T-n)}{(T-1)n}(\bar{\mu}-\mu)^T Q^{-1}(\bar{\mu}-\mu) \tag{3.5}$$

has a χ^2_n distribution with n degrees of freedom. If we replace Q with a positive definite estimate $\bar{Q}$, then the quantity in (3.5) has an F distribution with n and $T-n$ degrees of freedom; see Garlappi et al. [12] and Johnson and Wichern [17].

Garlappi et al. [12] consider the ellipsoidal uncertainty set

$$(\bar{\mu}-\mu)^T \bar{Q}^{-1}(\bar{\mu}-\mu) \leq \chi \ , \tag{3.6}$$

where $\chi = \frac{(T-1)n}{T(T-n)} q \geq 0$ and $q \geq 0$ is a chosen quantile for the distribution of (3.5), and simplify the min-max problem (3.3) subject to (3.6) and $e^T x = 1$ into the minimization problem

$$\begin{aligned} \min_x \quad & -\bar{\mu}^T x + \lambda x^T \bar{Q} x + \sqrt{\chi}\sqrt{x^T \bar{Q} x} \\ \text{s.t.} \quad & (\bar{\mu}-\mu)^T \bar{Q}^{-1}(\bar{\mu}-\mu) \leq \chi \\ & e^T x = 1 \ . \end{aligned} \tag{3.7}$$

Moreover, the optimal solution of (3.7) can be written as

$$x^* = \frac{1}{2\lambda}\bar{Q}^{-1}\left(\frac{1}{1+\frac{\sqrt{\chi}}{2\lambda\sigma_p^*}}\right)\left[\bar{\mu} - \frac{B - 2\lambda(1+\frac{\sqrt{\chi}}{2\lambda\sigma_p^*})}{A} e\right] \ , \tag{3.8}$$

where $A = e^T\bar{Q}^{-1}e$, $B = \bar{\mu}^T\bar{Q}^{-1}e$, and σ_p^* are the variances of the optimal portfolio which can be obtained from solving the polynomial equation

$$4A\lambda^2\sigma_p^4 + 4A\lambda\sqrt{\chi}\sigma_p^3 + (A\chi - AC + B^2 - 4\lambda^2)\sigma_p^2 - 4\lambda\sqrt{\chi}\sigma_p - \chi = 0\ . \qquad (3.9)$$

Given $\bar{Q}$ is positive definite, σ_p^* is the unique positive real root of (3.8). Note that since q corresponds to a confidence level with respect to the distribution of (3.5), χ can be interpreted as an estimation risk aversion parameter for the min-max robust optimization; the larger the value of χ, the more (estimation) risk aversion of the resulting min-max robust portfolios. As shown in Garlappi et al. [12], when $\chi = 0$ (either $T \to \infty$ or $q \to 0$), the estimation risk is ignored and the optimal portfolio (3.8) converges to the MV portfolio based on nominal estimates $\bar{\mu}$ and $\bar{Q}$. When $\chi \to \infty$, the optimal portfolio (3.8) converges to the minimum-variance portfolio which is generated without consideration of the portfolio expected return, i.e., the information of μ is not used for portfolio optimization.

To better understand the properties of the min-max robust portfolios computed from problem (3.7), we transform the objective function of (3.7) into

$$\min_x -\bar{\mu}^T x + \lambda(1 + \frac{\sqrt{\chi}}{\lambda\sqrt{x^T\bar{Q}x}})x^T\bar{Q}x\ . \qquad (3.10)$$

The formulation in (3.10) shows that, without the no-shortselling constraint $x \geq 0$, the min-max robust portfolio for the ellipsoidal uncertainty set (3.6) can be computed by solving the standard MV portfolio optimization problem (based on nominal estimates $\bar{\mu}$ and $\bar{Q}$) with a larger risk aversion parameter $\tilde{\lambda} = \lambda(1 + \frac{\sqrt{\chi}}{\lambda\sqrt{x^T\bar{Q}x}})$. In fact, the following theorem from Zhu et al. [33] shows that the solution of the problem

$$\begin{aligned} \min_{x} \quad & \max_{\mu} -\mu^T x + \lambda x^T \bar{Q} x \\ \text{s.t.} \quad & (\bar{\mu} - \mu)^T \bar{Q}^{-1} (\bar{\mu} - \mu) \leq \chi \\ & e^T x = 1, \quad x \geq 0 \end{aligned} \tag{3.11}$$

is also a solution of the problem

$$\begin{aligned} \min_{x} \quad & -\bar{\mu}^T x + \tilde{\lambda} x^T \bar{Q} x \\ & e^T x = 1, \quad x \geq 0 \end{aligned} \tag{3.12}$$

for some $\tilde{\lambda} \geq \lambda$, and the theorem holds regardless of whether the no-shortselling constraint $x \geq 0$ or additional linear constraints are imposed.

Theorem 3.1 (Zhu et al. [33, Theorem 2.2])**.** *Assume that $\bar{Q}$ is symmetric positive definite and $\chi \geq 0$. Any robust portfolio for the min-max robust mean-variance model (3.11) is an optimal portfolio for the standard mean-variance model based on nominal estimates $\bar{\mu}$ and $\bar{Q}$ with a risk aversion parameter $\tilde{\lambda} \geq \lambda$.*

Proof. Based on Theorem A.1 in Appendix A, the min-max robust MV problem

$$\begin{aligned} \min_{x} \quad & \max_{\mu} -\mu^T x + \lambda x^T \bar{Q} x \\ \text{s.t.} \quad & (\bar{\mu} - \mu)^T \bar{Q}^{-1} (\bar{\mu} - \mu) \leq \chi \\ & e^T x = 1, \quad x \geq 0 \end{aligned}$$

is equivalent to

$$\begin{aligned} & \min_{x} -\bar{\mu}^T x + \lambda x^T \bar{Q} x + \sqrt{\chi} \sqrt{x^T \bar{Q} x} \\ \text{s.t.} \quad & e^T x = 1, \quad x \geq 0 \,. \end{aligned}$$

Since this is a convex programming problem, it is easy to show that there exists $\tilde{\chi} \geq 0$ such that the above problem is equivalent to

$$\begin{aligned} &\min_x -\bar{\mu}^T x + \lambda x^T \bar{Q} x \\ \text{s.t.} \quad &\sqrt{x^T \bar{Q} x} \leq \tilde{\chi} \\ &e^T x = 1, \quad x \geq 0 \ . \end{aligned}$$

The above problem is equivalent to

$$\begin{aligned} &\min_x -\bar{\mu}^T x + \lambda x^T \bar{Q} x \\ \text{s.t.} \quad &x^T \bar{Q} x \leq \tilde{\chi}^2 \\ &e^T x = 1, \quad x \geq 0 \ . \end{aligned}$$

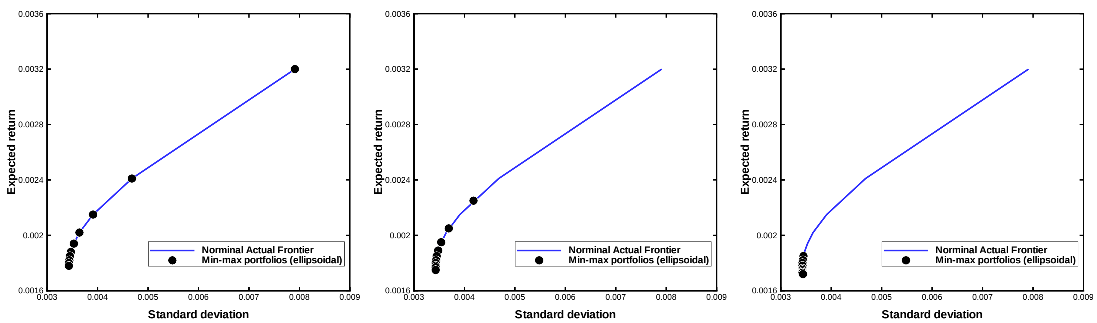
From the convexity of the problem and the Kuhn-Tucker conditions, there exists $\hat{\lambda} \geq 0$ such that the above problem is equivalent to

$$\begin{aligned} &\min_x -\bar{\mu}^T x + \lambda x^T \bar{Q} x + \hat{\lambda} x^T \bar{Q} x \\ \text{s.t.} \quad &e^T x = 1, \quad x \geq 0 \ . \end{aligned}$$

This completes the proof. □

Theorem 3.1 implies that the min-max robust MV model (3.11), which uses the ellipsoidal uncertainty set (3.6), adds robustness for estimation risk by increasing the risk aversion parameter λ. Thus the min-max robust actual frontiers from problem (3.11) are squeezed segments of the nominal actual frontiers from problem (2.9). Indeed, this can be illustrated by Figure 3.1(a)-(c) which compare the actual frontier segments of the two problems for different χ values used in (3.6). Each segment corresponds to the optimal portfolios for a sequence of λ values: $\lambda = [100, 200, 300, 400, 500, 600, 700, 800, 900, 1000]$. To differentiate the two actual frontier segments, we present the min-max one using a sequence of '$\bullet$' symbols, with each '$\bullet$' corresponds to the min-max robust portfolio for a particular λ value.

As we can see, regardless of the value of χ (which is interpreted as an estimation risk aversion parameter), all the min-max robust portfolios lie exactly on the actual frontier segment for the MV portfolio optimization problem (2.9). When $\chi = 0$, the actual frontier segments (as well as the resulting optimal portfolios correspond to the same λ) from problem (3.11) and problem (2.9) are identical. As the value of χ increases, the min-max actual frontier segment for problem (3.11) becomes shorter; this indicates the increase in estimation risk aversion.



(a) $\chi = 0$ (b) $\chi = 5$ (c) $\chi = 50$

Figure 3.1: Min-max robust portfolios (for the ellipsoidal uncertainty set (3.6)) and nominal actual frontier segment. Nominal actual frontiers are calculated from (2.9), which is the standard MV model that takes nominal estimates $\bar{\mu}$ and $\bar{Q}$ as input parameters. Min-max robust portfolios (with short-selling allowed) are computed from (3.8).

One potential problem of the min-max robust formulation (3.11) is that, based on Theorem 3.1, the performance of the resulting min-max robust portfolios depend on the accuracy of the point estimate $\bar{\mu}$. This means that a disastrous nominal actual frontier for problem (2.9) also leads to a disastrous min-max robust actual frontier for problem (3.11), as the later is a segment of the former. In the remaining parts of the thesis, instead of using an ellipsoidal structure for uncertainty sets, we focus on the min-max robust MV model that uses an interval structure for uncertainty sets.

3.2.2 Interval Uncertainty Set

Next, we illustrate characteristics of the min-max robust actual frontier for the min-max robust MV model (3.3) based on interval uncertainty sets. We consider using the interval structure presented in Tütüncü and Koenig [30] to represent uncertainty sets S_μ and S_Q, which are defined as:

$$\begin{aligned} S_\mu &= \{\mu : \mu^L \leq \mu \leq \mu^U\}, \\ S_Q &= \{Q : Q^L \leq Q \leq Q^U, Q \succeq 0\}, \end{aligned} \tag{3.13}$$

where μ^L, μ^U, Q^L and Q^U are the boundary values of these intervals. $Q \succeq 0$ indicates that Q is symmetric positive semi-definite, which is a necessary property for the unknown Q to be a covariance matrix of the returns. This property is a requirement for using the simple-case algorithm to compute min-max robust portfolios. As pointed by Tütüncü and Koenig [30], an interval structure can be flexible for defining the boundary values of the uncertainty sets. For example, the uncertainty set of μ, S_μ, can be obtained in the form of intervals by sampling μ from historical returns. Then the boundary values μ^L and μ^U can be defined as quantile values with respect to the generated μ sample distribution. By changing the quantile values for μ^L and μ^U, we can adjust the size and the bounds of S_μ. As a result, the performance of the resulting min-max robust portfolios may be changed. In this case, min-max robustness can be regarded as a quantile-based robustness approach.

Computing Min-max Robust Portfolios

To generate a min-max robust actual frontier, we must first discuss how the min-max robust portfolios on the frontier are computed. The min-max robust model (3.3) with interval uncertainty sets (3.13) was first introduced by Halldórsson and Tütüncü [15] as an application of using an interior-point method for solving saddle-point problems. Two algorithms for solving problem (3.3) are further discussed in Tütüncü and Koenig [30]: one is for the general case and the other is for the simple case. Here we discuss the algorithms for both cases.

The General Case

In general, problem (3.3) with the interval uncertainty sets defined in (3.13) are formulated by Halldórsson and Tütüncü [15] as the following saddle-point problem. Let $\phi(x, \mu, Q)$ denote the objective function in problem (3.3), i.e.,

$$\phi(x, \mu, Q) = -\mu^T x + \lambda x^T Q x, \tag{3.14}$$

where $\mu \in S_\mu$, $Q \in S_Q$, $x \in \Omega$ and $\lambda \geq 0$. When μ and Q are fixed, $\phi(x, \mu, Q)$ is a convex quadratic function of x; when x is fixed, $\phi(x, \mu, Q)$ is a linear function of μ and Q. Assuming that the feasible set Ω and the uncertainty sets S_μ and S_Q are all nonempty and bounded, Halldórsson and Tütüncü [15] show that the optimal value of the problem

$$\min_{x \in \Omega} \max_{(\mu \in S_\mu, Q \in S_Q)} \phi(x, \mu, Q) \tag{3.15}$$

is equal to that of the problem

$$\max_{(\mu \in S_\mu, Q \in S_Q)} \min_{x \in \Omega} \phi(x, \mu, Q). \tag{3.16}$$

This indicates that the function $\phi(x,\mu,Q)$ has a saddle-point, i.e, there exists a decision vector $\bar{x} \in \Omega$, and parameters $\bar{\mu} \in S_\mu$ and $\bar{Q} \in S_Q$ such that

$$\phi(x,\bar{\mu},\bar{Q}) \leq \phi(\bar{x},\bar{\mu},\bar{Q}) \leq \phi(\bar{x},\mu,Q) \tag{3.17}$$

Therefore, solving the min-max problem (3.3) is equivalent to finding a saddle-point of the function $\phi(x,\mu,Q)$, and an interior-point method is developed by Halldórsson and Tütüncü [15] for solving this saddle-point problem.

The Simple Case

In a special case , problem (3.3) can be solved as a standard QP problem. Tütüncü and Koenig [30] shows that, when $Q^U \succeq 0$, μ^L and Q^U are the optimal solutions for the problem:

$$\max_{(\mu \in S_\mu, Q \in S_Q)} -\mu^T x + \lambda x^T Q x \ , \quad \lambda \geq 0,$$

regardless of the values of (nonnegative) λ and vector x. In this case, the min-max problem (3.3) can be reduced to the following MV optimization problem:

$$\begin{aligned} \min_{x} \qquad & -(\mu^L)^T x + \lambda x^T (Q^U) x \\ \text{s.t.} \qquad & x \in \Omega \ . \end{aligned} \tag{3.18}$$

Therefore, we can determine a min-max robust portfolio by first finding μ^L and Q^U from their uncertainty intervals, and then solving problem (3.18) which takes μ^L and Q^U as the input parameters. Similarly, under the same assumptions, Tütüncü and Koenig [30] show that the problem (3.4), which is equivalent to the problem (3.3), can be reduced to the following MV optimization problem:

$$\begin{aligned} \min_x \quad & x^T(Q^U)x \\ \text{s.t.} \quad & (\mu^L)^T x \geq R \ , \\ & x \in \Omega \ . \end{aligned} \tag{3.19}$$

Generating Min-max Robust Actual Frontiers

To approximate the exact min-max robust actual frontier that contains infinite number of portfolio points, we can use the finite algorithm, as in Tütüncü and Koenig [30]. Let x_{min} and x_{max} denote the portfolios that achieve the minimum worst-case expected return, $R^{\text{w}}_{\text{min}}$, and the maximum worst-case expected return, $R^{\text{w}}_{\text{max}}$, respectively. This finite algorithm is presented as follows:

Algorithm 2 Generating Min-max Robust Actual Frontier

---

1. Compute x_{max} by solving problem (3.3) with $\lambda = 0$. Then set $R^{\text{w}}_{\text{max}} = (\mu^L)^T x_{\text{max}}$.
2. Compute x_{min} by solving problem (3.4) without the expected return constraint. Then set $R^{\text{w}}_{\text{min}} = (\mu^L)^T x_{\text{min}}$.
3. Generate m equally spaced values between $R^{\text{w}}_{\text{min}}$ and $R^{\text{w}}_{\text{max}}$ such that: $R^{\text{w}}_{\text{min}} \leq R^{\text{w}}_1 \leq R^{\text{w}}_2 \ldots \leq R^{\text{w}}_{\text{m}} \leq R^{\text{w}}_{\text{max}}$. For each $i \in m$, compute x_{i} by solving problem (3.4) with R^{w}_{i} as the worst-case expected return constraint.

Algorithm 2 is similar to Algorithm 1 (presented in Section 2.1.3) in the following sense: both algorithms generate the frontier by first determining the two extreme portfolios which have the maximum and the minimum expected returns, and then discretizing the range between these two extreme points with a finite number expected returns. As the number of discretization points becomes ∞, the approximated frontier approaches the exact frontier. However, unlike Algorithm 1 which is applied with the true expected return values and computes MV efficient portfolios by solving parametric QP problems, Algorithm 2 considers the expected returns in the worst-case scenario and computes min-max robust portfolios by solving parametric saddle-point problems. In addition, the min-max robust portfolios generated from Algorithm 2 may not be equally incremented when they are plotted with the true μ and Q. Note that for the simple case of the min-max robust MV model, the min-max robust portfolio in Step 1 can be computed by solving problem (3.18) instead of problem (3.3), and the ones in Step 2 and Step 3 can be computed by solving problem (3.19) instead of problem (3.4). In this case, Algorithm 1 and Algorithm 2 are conceptually identical.

3.3 Potential Problems of Min-max Robust MV Model

The min-max robust portfolio optimization model (3.3) provides a conservative framework to determine an optimal portfolio under the estimation risk of MV model parameters. It guarantees the portfolio performance under the worst sample scenarios of the uncertain parameters in their uncertainty sets. However, we will illustrate by the following example that this framework can be too sensitive to the initial data used to generate the uncertainty sets. In particular, inappropriate boundaries of interval uncertainty sets can cause min-max robust portfolios to be either too conservative or not conservative enough.

3.3.1 An Example

We illustrate these issues of min-max robust actual frontiers with a computational example. Comparisons are made with the true MV efficient frontier. We consider a universe of eight risky assets, and assume that their monthly mean vector μ and return covariance Q are given in Table 2.1. The true MV efficient frontier is generated by using the true μ and Q in Algorithm 1 (which is presented in Section 2.1.3). To compute the min-max robust actual frontier formed by min-max robust portfolios, we must first construct the uncertainty sets

S_μ and S_Q for μ and Q, respectively. Here we consider the situation where S_μ and S_Q are generated from the collections of discrete μ and Q samples.

Different sampling techniques can be used to generate μ and Q samples. For example, Tütüncü and Koenig [30] describe two methods with which the samples are generated based on historical data. The first one is based on bootstrapping, which repeatedly samples with replacement many series of k returns from the available observations. Then the μ and Q samples are computed from these series. The second one is based on moving averages, which considers moving windows of historical data, and computes the μ and Q samples in each window. Since both approaches are based on historical data, Tütüncü and Koenig [30] eliminate some of the lowest and highest quantiles of the resulting samples to minimize outlier effect.

RS Sampling Technique

For our computational example here, we construct the uncertainty sets by using the Monte Carlo re-sampling (RS) method introduced in Michaud [26]. RS is a statistical technique which constructs a distribution of the unknown data by sampling via Monte Carlo simulation from the original sample; a different sampling method based on the statistics (3.5) of sample mean estimates will be presented in Chapter 5.

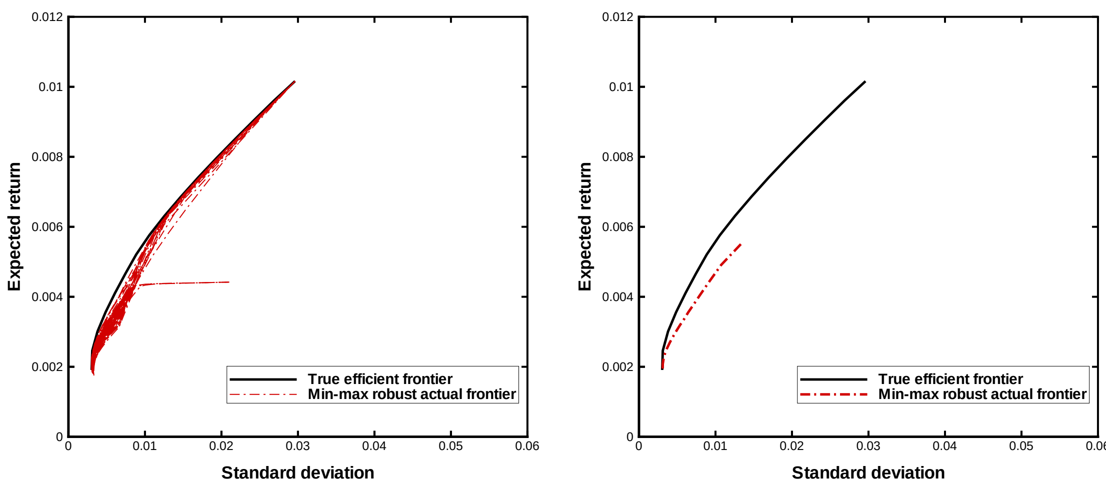
The process of sampling the mean with the RS technique is described as follows. Let's assume that 100 return samples are drawn from a multivariate normal distribution with mean μ and covariance matrix Q. We calculate the means and covariance from these 100 samples and denote them as $\bar{\mu}$ and $\bar{Q}$ respectively. Assume that $\bar{\mu}$ and $\bar{Q}$ are the representatives of μ and Q, we simultaneously generate 10,000 sets of independent return samples, each set consisting of 100 return samples. Regarding each set of 100 samples are equally likely to be observed, we compute the mean of each sample set and obtain these 10,000 estimates of mean returns as equally likely. These 10,000 estimates now form the uncertain universe for μ. One way of defining the boundary vectors μ^L and μ^U is to set them as the lowest and highest values respectively from these estimates. Tütüncü and Koenig [30] show that Q^L and Q^U can be determined by selecting the lowest and highest values of the covariance between each pair of assets (and the variance of each asset) respectively. However, as demonstrated by Broadie [7], the impact of the estimation error on μ is much larger, and the estimation for Q is much more accurate. So here we ignore the estimation error on Q, and set both Q^L and Q^U equal to $\bar{Q}$. Since $\bar{Q}$ obtained for this example is positive semi-definite, we can compute min-max robust portfolios by solving the simple case problem (3.18).

Sensitivity to Initial Return Samples

To illustrate the sensitivity of min-max robust portfolios to the choice of initial return samples used to generate S_μ, the above μ sampling process is repeated 100 times, with different set of 100 return samples at each time. All the min-max robust actual frontiers are graphed, together with the true efficient frontier, in Figure 3.2(a). For this example, min-max robust actual frontiers have relatively small variation near their minimum-variance portfolios, and as the expected return increases, the variation becomes larger. A few of the actual frontiers are very close to the true efficient frontier and achieve the same maximum expected return. To further assess the frontiers, we depict the average performance: for the portfolios on the min-max robust actual frontiers in Figure 3.2(a) corresponding to the same λ value, we calculate their average returns and variances. Doing this calculation for all λ values, we obtain the "average" min-max robust actual frontier presented in Figure 3.2(b).

Over-Conservative Problem

Figure 3.2(b) illustrates that min-max robust portfolios can be too conservative; the maximum expected returns of min-max robust actual frontiers can be significantly lower. Indeed, the maximum expected return of the true efficient frontier is 1.02×10^{-2}, while the corresponding value of the "average" min-max robust actual frontier is only 0.56×10^{-2}. Min-max robust optimization provides an effective approach for conservative investors to prevent undesirable losses due to the estimation risk in MV model parameters. But it may not be an appropriate choice for the investors who are more tolerant to estimation risk and wish to seek higher returns: their expected returns may not be achievable by min-max robust portfolios.

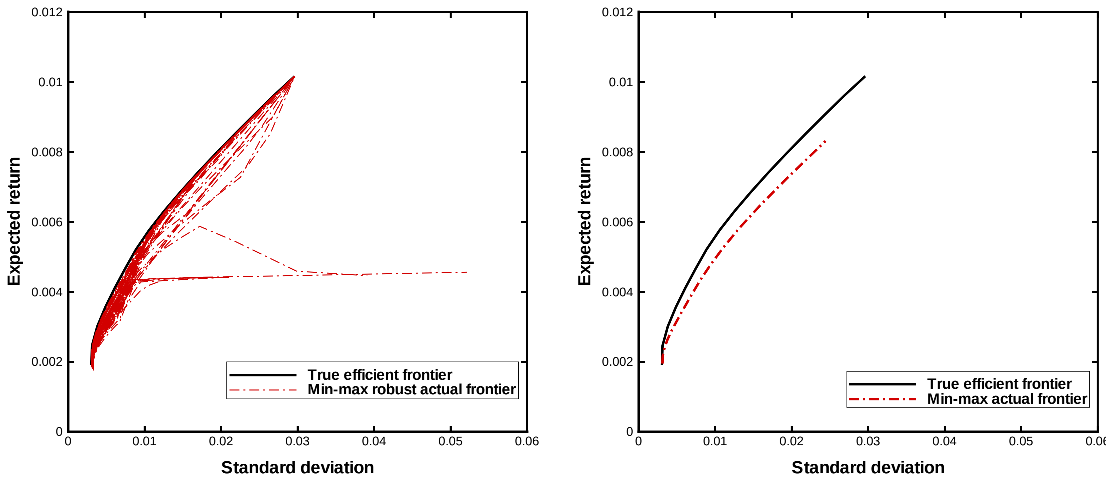


(a) 100 min-max robust actual frontiers

(b) "Average" min-max robust actual frontier

Figure 3.2: Min-max actual frontiers for the 8-asset example in Table 2.1.

It is true that the conservatism of the min-max robust portfolios generated from the above example can be adjusted by changing the uncertainty interval. For example, Figure 3.3(a) and 3.3(b) illustrate the min-max robust actual frontiers and their "average" actual frontier obtained based on using 2.5 and 97.5 percentiles of S_μ as the values for μ^L and μ^U. However, it is important to note that the worst sample scenario is eliminated in the new uncertainty set and no longer plays any role in the resulting robust portfolio. Thus in the min-max robust MV model, the conservatism is adjusted by excluding bad scenarios, which may run counter to the principal of min-max robustness.



(a) 100 min-max robust actual frontiers (b) "Average" min-max robust actual frontier

Figure 3.3: Min-max robust actual frontiers (with improved μ^L) for the 8-asset example in Table 2.1.

Compared with the min-max robust actual frontiers in Figure 3.2, the frontiers in Figure 3.3 become longer, and achieve a higher maximum expected return in the "average" case. This is due to the fact that the worst 2.5% μ-samples are excluded (and the length of the uncertainty interval S_μ becomes shorter). It also shows that the min-max robust MV model is very sensitive to the choice of the boundary μ^L. Therefore, the uncertainty set must be carefully chosen. In particular, inappropriate μ^L can make the resulting min-max robust portfolios either too conservative or not conservative enough.

3.4 Robustness of Actual Frontiers

One way of assessing the robustness of a portfolio optimization technique is to measure the variation of its actual frontiers calculated from different estimated parameters. Less variation in the actual frontiers may be considered more robust. Comparing with the min-max robust actual frontiers in Figure 3.2(a), the frontiers in Figure 3.3(a) exhibit more variation. Indeed, min-max robust portfolios can be quite sensitive to the initial return samples used to generate the uncertainty set. For additional illustration, we use the example in Best and Grauer [5], where the portfolio is constructed using ten assets. The true mean vector and covariances of returns are given in Table 3.1. We conduct the same computation as in Figure 3.2, and plot the results in Figure 3.4. As can be observed, the min-max robust actual frontiers have large variations not only near the maximum-return portfolios but also near the minimum-variance portfolios.

Table 3.1 Mean vector and covariance matrix for a 10-asset portfolio problem

Mean Return Vector μ

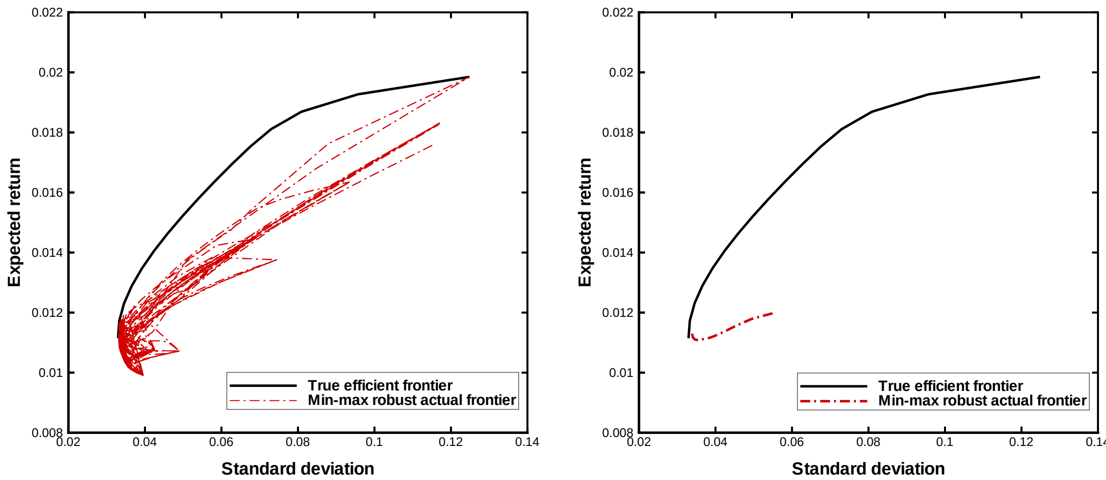
$10^{-2}\times$	Asset1	Asset2	Asset3	Asset4	Asset5	Asset6	Asset7	Asset8	Asset9	Asset10
	1.0720	1.7618	1.8270	1.0761	1.9845	1.4452	0.9910	1.6353	1.3755	1.8315

Covariance Matrix Q

$10^{-2}\times$	Asset1	Asset2	Asset3	Asset4	Asset5	Asset6	Asset7	Asset8	Asset9	Asset10
Asset1	0.2516									
Asset2	0.0766	1.3743								
Asset3	0.1104	0.2847	1.3996							
Asset4	0.1314	0.0930	0.1027	0.1928						
Asset5	0.0157	0.5610	0.4725	0.0451	1.5981					
Asset6	0.0554	0.3457	0.2769	0.0898	0.3490	0.4787				
Asset7	0.0937	0.0253	0.0759	0.1010	0.0714	0.0643	0.1664			
Asset8	0.1646	0.1757	0.3200	0.1641	0.4721	0.2669	0.1020	0.9013		
Asset9	0.0509	0.1810	0.3275	0.0993	0.2978	0.1783	0.0635	0.1534	0.5731	
Asset10	0.1515	0.3445	0.3627	0.0966	0.4740	0.2651	0.0611	0.3596	0.2154	1.4041

3.4.1 Variance-based Actual Frontiers

If the sole criterion for a robust approach is the least variation of actual frontiers with respect to the initial data, we may consider an alternative approach. Relative to the mean return μ, the covariance matrix Q can be estimated more accurately. Thus we can generate an actual frontier based only on the estimation of Q as follows.



(a) 100 min-max robust actual frontiers

(b) "Average" min-max robust actual frontier

Figure 3.4: Min-max robust actual frontiers for the 10-asset example in Table 3.1 .

Assume that $\bar{Q}$ is the estimated covariance matrix of the normally distributed asset returns, and let $x_{\min}$ and $x_{\max}$ denote the portfolios that have the minimum and maximum variance, respectively, i.e., $x_{\min}$ and $x_{\max}$ are the respective optimal solutions for the two QP problems:

$$\begin{aligned} &\min_x x^T \bar{Q} x \\ &\text{s.t. } \ x \in \Omega \end{aligned} \tag{3.20}$$

$$\begin{aligned} &\max_x x^T \bar{Q} x \\ &\text{s.t. } \ x \in \Omega \ . \end{aligned} \tag{3.21}$$

For any parameter $\eta \in [0, 1]$, we can compute a portfolio $x(\eta) \in R^n$ as below:

$$x(\eta) = (1 - \eta)x_{\min} + \eta x_{\max}. \tag{3.22}$$

Given that both $x_{\min}$ and $x_{\max}$ are in the feasible set Ω, it is easy to show that all $x(\eta),\ 0 \leq \eta \leq 1$, belong to Ω.

Proposition 1. *All portfolios $x \in R^n$ computed from (3.22) belong to the convex feasible set $\Omega = \{x \in \mathbb{R}^n \mid \sum_{i=1}^n x_i = 1, x \geq 0\}$.*

Proof. First, note that since both $x_{\min}$ and $x_{\max}$ are non-negative and $\eta \in [0, 1]$, any x computed from equation (3.22) will be non-negative. Also, since $\sum_{i=1}^n (x_{\min})_i = 1$ and $\sum_{i=1}^n (x_{\max})_i = 1$, we have:

$$\begin{aligned}
\sum_{i=1}^n x_i &= \sum_{i=1}^n (1 - \eta)(x_{\min})_i + \sum_{i=1}^n \eta(x_{\max})_i \\
&= (1 - \eta)\sum_{i=1}^n (x_{\min})_i + \eta \sum_{i=1}^n (x_{\max})_i \\
&= (1 - \eta) + \eta \\
&= 1.
\end{aligned}$$

Therefore, $x \in \Omega$. □

We compute 100 such actual frontiers based on 100 estimates of the true covariance matrix Q in Table 3.1, and plot them in Figure 3.5(a). Each covariance matrix estimate is based on 100 return samples. Compared with the min-max robust actual frontiers in Figure 3.4(a), the variance-based actual frontiers in Figure 3.5(a) have some attractive properties as far as robustness is concerned. Firstly, the variance-based actual frontiers have much smaller variations. Secondly, min-max robust portfolios achieving higher expected returns tend to have larger fluctuations. On the other hand, actual frontiers of the variance-based portfolios maintain similar variations for different risks. Thirdly, for this example, the variation of variance-based actual frontiers decreases as the risk increases. If a resulting actual frontier has a sufficiently large and positive slope as in Figure 3.5(a), then this variance-based approach could be effective for obtaining a portfolio which not only achieves a high expected return but also is relatively robust in the sense that its composition based on different estimations of Q have small variations. The following theorem provides a condition which guarantees monotonicity of the actual frontier from the variance-based approach.

Theorem 3.2. *Suppose Q is symmetric positive semi-definite and η is a real number such that $0 \leq \eta \leq 1$. Let x_{min}, x_{max} and $x(\eta)$ be the values computed from problem (3.20), (3.21) and equation (3.22), respectively. (a) If $\mu^T x_{max} \geq \mu^T x_{min}$, then $\mu^T x(\eta_1) \leq \mu^T x(\eta_2)$ for any $0 \leq \eta_1 \leq \eta_2 \leq 1$. (b) If $x_{min}^T Q x_{max} \geq x_{min}^T Q x_{min}$, then $f(\eta) = x(\eta)^T Q x(\eta)$ is a non-decreasing function of η, i.e., $f(\eta_1) \leq f(\eta_2)$ for any $0 \leq \eta_1 \leq \eta_2 \leq 1$.*

Proof. From $\mu^T x(\eta) = (1-\eta)\mu^T x_{min} + \eta\mu^T x_{max}$, it immediately follows that

$$\mu^T x(\eta_2) - \mu^T x(\eta_1) = (\eta_2 - \eta_1)(\mu^T x_{max} - \mu^T x_{min}) \geq 0, \quad \text{for any } 0 \leq \eta_1 \leq \eta_2 \leq 1. \tag{3.23}$$

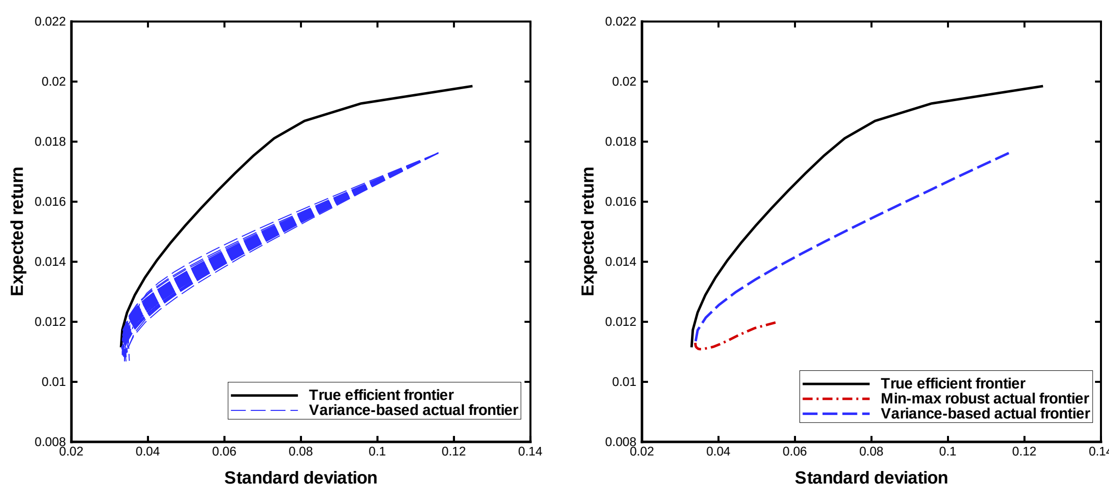
To prove $f(\eta) = x(\eta)^T Q x(\eta)$ is a non-decreasing function of η when $x_{\min}^T Q x_{\max} \geq x_{\min}^T Q x_{\min}$, it is sufficient to prove that $f'(\eta) \geq 0$ under the same condition.

Given $x(\eta) = (1-\eta)x_{\min} + \eta x_{\max}$, we have:

$$\begin{aligned} f(\eta) &= x(\eta)^T Q x(\eta) \\ &= ((1-\eta)x_{\min} + \eta x_{\max})^T Q((1-\eta)x_{\min} + \eta x_{\max}) \\ &= (1-\eta)^2 x_{\min}^T Q x_{\min} + 2\eta(1-\eta) x_{\min}^T Q x_{\max} + \eta^2 x_{\max}^T Q x_{\max} \\ f'(\eta) &= 2(\eta-1) x_{\min}^T Q x_{\min} + (2-4\eta) x_{\min}^T Q x_{\max} + 2\eta x_{\max}^T Q x_{\max} \\ &= 2\eta x_{\min}^T Q x_{\min} - 4\eta x_{\min}^T Q x_{\max} + 2\eta x_{\max}^T Q x_{\max} + 2x_{\min}^T Q x_{\max} - 2x_{\min}^T Q x_{\min} \\ &= 2\eta(x_{\min} - x_{\max})^T Q(x_{\min} - x_{\max}) + 2x_{\min}^T Q(x_{\max} - x_{\min}) \end{aligned}$$

Since $x_{\min}^T Q x_{\max} \geq x_{\min}^T Q x_{\min}$, $x_{\min}^T Q(x_{\max} - x_{\min}) \geq 0$. Therefore, $f'(\eta) \geq 0$ for any $0 \leq \eta \leq 1$. Thus $f(\eta)$ is a non-decreasing function of η.

This completes the proof. □

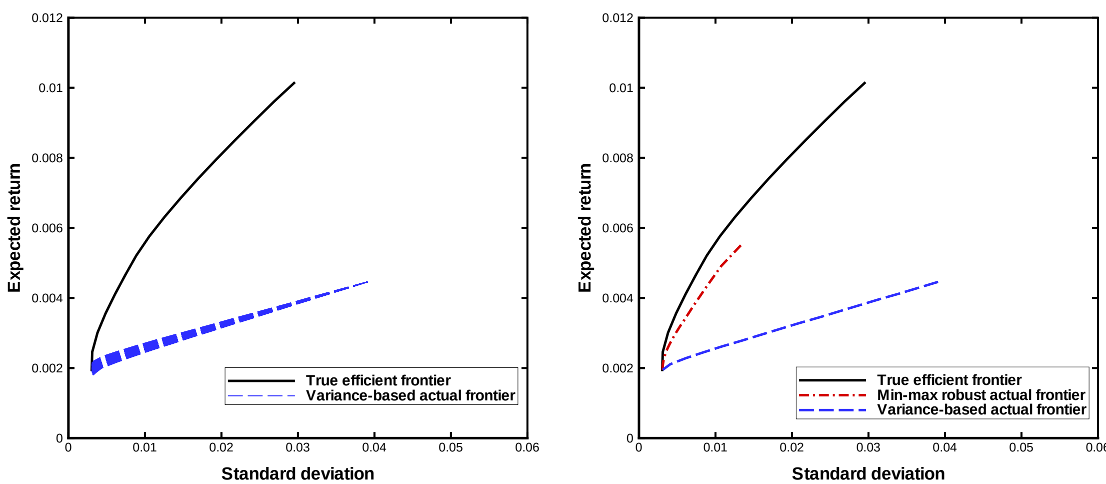


(a) 100 variance-based actual frontiers (b) "Average" variance-based actual frontier

Figure 3.5: Variance-based actual frontiers for the 10-asset example in Table 3.1 .

3.4.2 Beyond Variation

It is interesting to investigate how the variance-based approach (3.22) compares with the min-max robust approach (3.3) in terms of portfolio efficiency. One approach is more efficient than another if its resulting portfolio achieves a higher expected return for the same level of risk, or achieves a lower risk for the same level of expected return. We compute the "average" variance-based actual frontiers in Figure 3.5(a) and plot it in Figure 3.5(b). The "average" min-max robust actual frontier from Figure 3.4(b) is also plotted in the same graph. By comparing the two "average" actual frontiers in Figure 3.5(b), we observe that the variance-based portfolios are more efficient for this 10-asset example. However, we demonstrate through another example that sometimes min-max robust portfolios achieve better performance.



(a) 100 variance-based actual frontiers

(b) "Average" variance-based actual frontier

Figure 3.6: Variance-based actual frontiers for the 8-asset sample in Table 2.1 .

Following the same procedure, we compute the variance-based actual frontiers for the 8-asset example in Table 2.1. The actual frontiers and their "average" are plotted in Figure 3.6(a) and (b), respectively. Similar to the 10-asset example, the variance-based portfolios in Figure 3.6(a) are more stable than the min-max robust portfolios in Figure 3.2(a). However, the min-max robust portfolios are more efficient in the "average" case. As shown in Figure 3.6(b), the "average" min-max robust actual frontier lies above the variance-based one. In Figure 3.6(b), the maximum expected return of the "average" min-max robust actual frontier is 0.56×10^{-2}, while the corresponding value of the variance-based one is only 0.45×10^{-2}. Thus the performance of the portfolios generated from the variance-based model can be very poor. This indicates that it may not be appropriate to evaluate the effectiveness of a robust model for MV portfolio optimization solely based on the variations of the resulting actual frontiers; the efficiency of the resulting actual frontiers is also important.

3.5 Conclusion and Remarks

In this chapter, we focus on the min-max robust MV model (with interval uncertainty sets) and show how to compute min-max robust portfolios and actual frontiers using this model. Through computational examples, we illustrate that the solution to the min-max robust MV model can be very sensitive to the initial data which is used to construct uncertainty sets. In addition, min-max robust portfolios can be too conservative to achieve sufficiently high expected return. We also demonstrate through examples that variation in the actual frontiers should not be the only criterion for evaluating the performance of a robust portfolio optimization model, otherwise the presented variance-based model can be much more effective. Eliminating poor mean samples adjusts the performance of min-max robust portfolios. However, doing this may face the risk of loosing the protection against the worst sample scenario, and is against the objective of robust optimization. In next chapter, we present a CVaR robust MV portfolio optimization model which adjusts the conservatism of robust portfolios without eliminating the worst mean samples.

Chapter 4

CVaR Robust Mean-Variance Portfolio Optimization

In Chapter 3, we have illustrated that the min-max robust mean-variance portfolio optimization model (3.3) can be very sensitive to the initial data from which the uncertainty set is specified. In addition, the robust portfolios produced by the min-max model can be too conservative for an investor who wants to obtain a higher portfolio return by taking more estimation risk in model parameters. For the min-max robust MV model, adjusting the uncertainty set to generate a less conservative portfolio is achieved by eliminating the worst sample scenarios, which runs counter to the robust objective. Moreover, the min-max robust optimization, by definition, neglects any probability information on the mean return distribution once the uncertainty set is specified.

In this chapter, we focus on the uncertainty of mean return, and propose a CVaR robust MV portfolio optimization model in which the estimation risk in mean return is measured by the Conditional Value-at-Risk (CVaR). In contrast to the min-max robust MV model, which takes the worst sample scenario in the uncertainty set of mean return, the CVaR robust MV model determines an optimal portfolio based on a tail of the mean loss distribution due to mean return uncertainty. The adjustment of the CVaR's confidence level according to investors' preferences corresponds to the adjustment of the conservatism level with respect to mean return uncertainty.

This chapter is divided into four sections. In Section 4.1, we summarize the definition of CVaR for a general loss distribution. In Section 4.2, we describe how CVaR is used to measure portfolio return risk in the traditional return-risk analysis. In Section 4.3, we use CVaR to measure the estimation risk in mean return. In Section 4.4, we present our CVaR robust MV portfolio optimization model and discuss its attractive properties as compared with min-max robust MV model.

4.1 CVaR for a General Loss Distribution

This section introduces the general definition of the Conditional Value-at-Risk (CVaR), which is based on another popular risk measure, the Value-at-Risk (VaR). VaR is widely used by financial institutions to quantify the market risk of portfolios. As defined in Rockafellar and Uryasev [29], the VaR of a portfolio is the lowest amount α such that, with a given probability β over a certain time period T, the portfolio loss will not exceed α. In other words, VaR is simply a number that indicates the worst-case loss of a portfolio with provability β over a certain time period T. CVaR, which is also known as Mean Shortfall or Mean Excess Loss, is defined as the conditional expectation of the losses exceeding VaR. Compared with VaR, CVaR is a coherent risk measure and has more attractive properties such as sub-additivity and convexity; see e.g., Artzner et al. [2], Rockafellar and Uryasev [29].

Consider a specific risk denoted by a random variable L (which typically corresponds to loss). Assume that L has a density function $p(l)$. The probability of L not exceeding a threshold α is represented as:

$$\Psi(\alpha) = \int_{l \leq \alpha} p(l)\, dl. \tag{4.1}$$

Under the assumption that the probability distribution of L has no jumps, $\Psi(\alpha)$ is everywhere continuous with respect to α.

Given a confidence level $\beta \in (0,1)$, e.g., $\beta = 95\%$, the associated Value-at-Risk, VaR_β, is defined as:

$$\mathrm{VaR}_\beta = \min \{\alpha \in R : \Psi(\alpha) \geq \beta\} \ . \tag{4.2}$$

The corresponding CVaR, denoted by CVaR_β, is given by

$$\mathrm{CVaR}_\beta = \mathbb{E}(L \mid L \geq \mathrm{VaR}_\beta) = \frac{1}{1-\beta} \int_{l \geq \mathrm{VaR}_\beta} l p(l) dl \tag{4.3}$$

when the loss distribution has no jumps. Thus, CVaR_β is the expected loss conditional on the loss being greater than or equal to VaR_β with probability $(1-\beta)$.

In addition, CVaR has the following equivalent expression,

$$\mathrm{CVaR}_\beta = \min_\alpha (\alpha + (1-\beta)^{-1} \mathbb{E}([L-\alpha]^+)) \ , \tag{4.4}$$

where $[z]^+ = \max(z, 0)$, see Rockafellar and Uryasev [29]. Note that, while VaR is a quantile, CVaR depends on the entire tail of the worst scenarios corresponding to a given confidence level.

4.2 A Traditional Measure for the Portfolio Return Risk

In the traditional return-risk analysis, the portfolio's expected return is the weighted combination of the assets' mean returns, and the associated risk is the volatility of portfolio return. Assuming that an investor is risk-averse, there exists a trade-off between portfolio expected return and risk, i.e., an investor will not get a higher expected return without taking a higher risk. Therefore, a portfolio is considered to be optimal (MV efficient) if it has the lowest risk for a given level of expected return, or conversely, has the maximum expected return for a given level of risk. When portfolio return is not normally distributed, CVaR can be used as an alternative to variance to measure the risk due to return volatility. To emphasize, this CVaR risk measure is denoted as CVaR^r, and its associated VaR risk measure is denoted as VaR^r.

Let $r \in R^n$ be the vector of the asset returns of n risky assets. We assume that the returns r constitute a joint normal distribution with the density $p(r)$. Let the decision vector $x \in \Omega$ be the portfolio weights, and denote $f(x, r)$ as the portfolio loss function associated with a fixed x. The portfolio return loss $f(x, r)$ is the negative of the portfolio return:

$$f(x, r) = -r^T x = -[r_1 x_1 + r_2 x_2 + \ldots + r_n x_n]. \tag{4.5}$$

Therefore, the CVaR_β of the return loss $f(x, r)$, $\text{CVaR}^r_\beta(x)$, can be defined by replacing the general loss L with $f(x, r)$ in formula (4.4):

$$\text{CVaR}^r_\beta(x) = \min_\alpha(\alpha + (1 - \beta)^{-1}\mathbb{E}([f(x, r) - \alpha]^+)). \tag{4.6}$$

Typically, one can consider portfolio optimization using CVaR risk measure rather than variance risk measure. Rockafellar and Uryasev [29] introduce a linear programming (LP) approach to solve the CVaR minimization problem, and apply this approach to portfolio optimization where optimal portfolios are computed by minimizing CVaR^r under a constraint on the expected return. By specifying different levels of expected returns, a sequence of optimal portfolios can be generated ranging from the minimum return risk (measured by CVaR^r) to the maximum expected return. Krokhmal et al. [19] extend this CVaR optimization technique, and show that an equivalent efficient frontier can be generated by maximizing the expected return under CVaR^r constraints. Imposing various CVaR^r constraints with different confidence levels, the portfolio loss distribution can be shaped according to different preferences of decision makers.

VaR^r and CVaR^r have also been used in robust portfolio optimization, and many efficient techniques have been developed to solve the associated robust optimization problems. Goldfarb and Iyengar [14] study robust VaR^r portfolio selection problem, where the objective is to maximize the worst-case expected return subject to the constraint that the shortfall probability is less than a prescribed limit. They show that under the normality assumption of asset return distribution, the robust optimization problem can be cast as a second-order cone programming (SOCP) problem. Zhu and Fukushima [34] optimize portfolios under the worst-case CVaR^r constraint in the situation where the information on the underlying return distribution is not exactly known but belongs to a certain set of distributions. Then the worst-case CVaR^r is the CVaR^r with respect to the worst-case scenario from the distribution set. They show that when the asset returns have a discrete distribution, using box uncertainty and ellipsoidal uncertainty structures, the robust optimization problem can be solved efficiently as LP problems and SOCP problems, respectively.

4.3 CVaR for the Estimation Risk in Mean Return

The CVaR risk measure discussed in Section 4.2, CVaR^r, is used to measure the portfolio return risk in the context of return-risk analysis. However, here our concern is the uncertainty of mean return when using the MV portfolio optimization model. In this section, we choose CVaR to be the measure of the estimation risk in mean return in the MV portfolio optimization. Applying CVaR with this perspective, we introduce a CVaR robust MV portfolio optimization model in Section 4.4 which computes optimal portfolios under mean return uncertainty. We use CVaR^μ to denote the CVaR with respect to mean return in order to differentiate it from CVaR^r with respect to return, and consider the portfolio loss quantified by CVaR^μ as the mean loss. Similarly, we denote VaR^μ as the VaR of mean loss. Different from CVaR^r, which quantifies the portfolio loss due to return volatility, CVaR^μ quantifies the portfolio mean loss due to mean return uncertainty.

For a portfolio of n risky assets, we let the decision vector $x \in \Omega$ be the portfolio weights, and $\mu \in R^n$ be the random vector of the asset mean returns. We assume that μ has a probability density function $p(\mu)$, and is independent of x. To determine the portfolio mean loss, we define $f(x, \mu)$, which is the mean loss function associated with μ, to be the negative of the portfolio expected return function:

$$f(x, \mu) = -\mu^T x = -[\mu_1 x_1 + \ldots + \mu_n x_n]. \tag{4.7}$$

Assume that μ is unknown but has a certain distribution, the mean loss $f(x, \mu)$ associated with all possible μ will also form a distribution. Therefore, as depicted in Figure 4.1, CVaR^μ_β corresponds to the mean of the $(1-\beta)$-tail, which contains the $(1-\beta)$ worst-case mean losses due to μ uncertainty. The definition of CVaR^μ_β can be derived from its general form in (4.4) by replacing the general loss L with the mean loss $f(x, \mu)$:

$$\text{CVaR}^\mu_\beta(x) = \min_\alpha (\alpha + (1-\beta)^{-1}\mathbb{E}([f(x, \mu) - \alpha]^+)). \tag{4.8}$$

4.4 CVaR Robust MV Portfolio Optimization Model

4.4.1 Model Definition

The traditional MV portfolio optimization model (2.6) computes an optimal portfolio by minimizing the portfolio mean loss $-\mu^T x$ (or maximizing the portfolio expected return $\mu^T x$) under the assumption that μ is the true value. Given that μ is unknown in practice, we

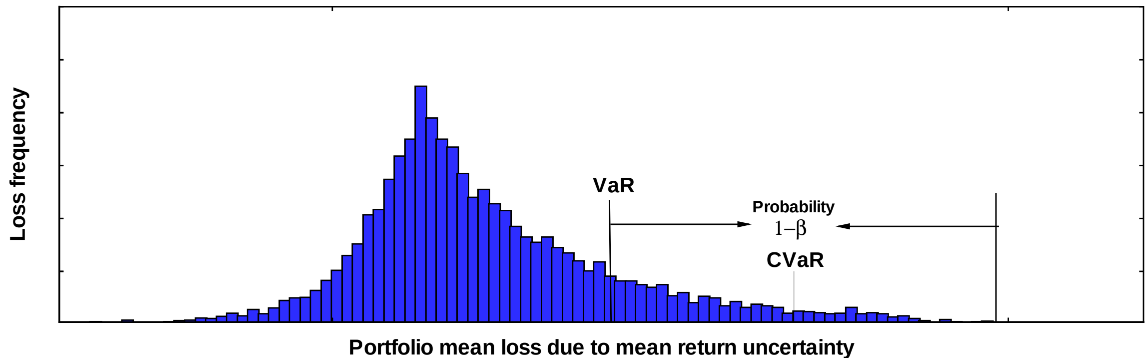


Figure 4.1: CVaR^{μ} for a portfolio mean loss distribution

consider the average of a tail of mean loss scenarios instead of a single mean loss. This takes the estimation error in μ into account and is achieved by replacing the mean loss $-\mu^T x$ by $\text{CVaR}^{\mu}_{\beta}(-\mu^T x)$ in the MV model. Therefore, an optimal CVaR robust portfolio is determined as the solution of the following optimization problem:

$$\begin{aligned} \min_{x} \quad & \text{CVaR}^{\mu}_{\beta}(-\mu^T x) + \lambda x^T \bar{Q} x \\ \text{s.t.} \quad & x \in \Omega, \end{aligned} \tag{4.9}$$

where $\bar{Q}$ is an estimate of the covariance matrix Q. Recall that we ignore in this thesis the estimation risk in Q as it is much smaller than that in μ. Problem (4.9) is the general formulation of our CVaR robust mean-variance (MV) portfolio optimization model. It determines the optimal portfolio by minimizing the risk-adjusted $\text{CVaR}^{\mu}_{\beta}$, which is the sum of $\text{CVaR}^{\mu}_{\beta}(-\mu^T x)$ (the CVaR of mean loss for a confidence level β) and $\lambda x^T \bar{Q} x$ (the portfolio variance for a specific risk-aversion parameter λ). Note that we consider minimizing the CVaR^{μ} with respect to an assumed distribution of μ. When the distribution only contains the true μ, i.e., no deviation in μ, the CVaR robust MV model is equivalent to the traditional MV model.

As discussed in Chapter 3, the min-max robust MV model computes robust portfolios based on the worst-case scenario μ^L when the uncertainty set of μ is the interval $[\mu^L,\, \mu^U]$. To determine a robust portfolio, the task is then switched to determine μ^L. Thus the resulting min-max robust portfolio can become sensitive to the initial data used to construct the uncertainty interval. In addition, the min-max approach is sometimes over conservative for an investor who wants to get a higher portfolio return by taking more estimation risk.

In contrast to the min-max robust MV model (3.3), the CVaR robust MV model (4.9) depends on the entire $(1-\beta)$-tail of the mean loss distribution. Using the CVaR robust MV model, adjusting the confidence level β of CVaR^μ naturally corresponds to adjusting an investor's tolerance to estimation risk. In general, when the β value increases, the CVaR^μ_β of the mean loss distribution will increase. This corresponds to the situation where the model takes more pessimistic view on the estimation risk in μ, and optimizes a portfolio under worse cases of the mean loss. Therefore, the resulting CVaR robust portfolio is forced to be more conservative. As $\beta \to 1$, the worst mean return sample is considered, and thus the model provides the most conservative portfolio. Following the same reason, when the β value decreases, less emphasis is placed on worse mean loss scenarios, and the resulting CVaR robust portfolio becomes less conservative. As $\beta \to 0$, all cases of the mean loss (ranging from its best case to its worst case) are considered; this leads to the least robustness in the sense of protection against the worst case. Note that the choice of β (or portfolio's robustness) implicitly affects the portfolio's expected return: the maximum expected return achievable for a higher β is generally less than that for a lower β. The choice of β depends on an individual investor's risk-averse preference with respect to the estimation risk in μ. We will demonstrate the impact of the β value on a portfolio's conservatism level through computational examples in Chapter 5.

4.4.2 Computing CVaR Robust Portfolios

Problem (4.9) can be solved by utilizing the CVaR optimization approach introduced by Rockafellar and Uryasev [29]. Define the following auxiliary function

$$F_\beta^\mu(x,\alpha) = \alpha + \frac{1}{1-\beta}\int_{\mu\in R^n} [f(x,\mu)-\alpha]^+ p(\mu)\, d\mu \; . \tag{4.10}$$

Assume that the distribution for μ is continuous, CVaR_β^μ is convex with respect to x, and $F_\beta^\mu(x,\alpha)$ is both convex and continuously differentiable. Therefore, Rockafellar and Uryasev [29] show that, for any fixed $x \in \Omega$, CVaR_β^μ can be determined by

$$\mathrm{CVaR}_\beta^\mu(x) = \min_\alpha F_\beta^\mu(x,\alpha) \; . \tag{4.11}$$

In addition, minimizing $\mathrm{CVaR}_\beta^\mu(x)$ over x is equivalent to minimizing $F_\beta^\mu(x,\alpha)$ over (x,α), i.e.,

$$\min_x \mathrm{CVaR}_\beta(x) \equiv \min_{x,\alpha} F_\beta^\mu(x,\alpha) \; . \tag{4.12}$$

Similarly, it can be shown that minimizing $\mathrm{CVaR}_\beta^\mu(-\mu^T x) + \lambda x^T \bar{Q} x$ over x is equivalent to minimizing $F_\beta^\mu(x,\alpha) + \lambda x^T \bar{Q} x$ over (x,α).

Theorem 4.1. *Given that $CVaR_\beta^\mu(-\mu^T x)$ is convex with respect to x, and $F_\beta^\mu(x,\alpha)$ is convex with respect to (x,α), the two minimization problems below*

$$\begin{array}{ll} \min\limits_x & CVaR_\beta^\mu(-\mu^T x) + \lambda x^T \bar{Q} x \\ & x \in \Omega \end{array} \tag{4.13}$$

$$\begin{array}{ll} \min\limits_{x,\alpha} & F_\beta^\mu(x,\alpha) + \lambda x^T \bar{Q} x \\ & x \in \Omega \end{array} \tag{4.14}$$

are equivalent for $\lambda \ge 0$.

Proof. Since the term $\lambda x^T \bar{Q} x$ is convex with respect to x, $\mathrm{CVaR}_\beta^\mu(x) + \lambda x^T \bar{Q} x$ is convex with respect to x and $F_\beta^\mu(x,\alpha) + \lambda x^T \bar{Q} x$ is convex with respect to (x,α). Therefore, we can minimize $F_\beta^\mu(x,\alpha) + \lambda x^T \bar{Q} x$ over $(x,\alpha) \in \Omega \times \mathbb{R}$ by first minimizing it over $\alpha \in \mathbb{R}$ for a fixed x and then minimizing the result over $x \in \Omega$. Moreover, as a consequence of (4.11), we have :

$$\mathrm{CVaR}_\beta^\mu(x) + \lambda x^T \bar{Q} x = \min_\alpha F_\beta^\mu(x,\alpha) + \lambda x^T \bar{Q} x \ . \tag{4.15}$$

Thus, the equivalence between problem (4.13) and (4.14) is established by minimizing both sides of (4.15) over $x \in \Omega$. □

We can approximate the integral in (4.10) by sampling μ based on its density function $p(\mu)$. For a collection of m independent samples $\mu_1, \mu_2, \ldots, \mu_m$, the corresponding approximation to function $F_\beta^\mu(x,\alpha)$ is:

$$\bar{F}_\beta^\mu(x,\alpha) = \alpha + \frac{1}{m(1-\beta)} \sum_{i=1}^{m} [-\mu_i^T x - \alpha]^+ \ . \tag{4.16}$$

Clearly, function $\bar{F}_\beta^\mu(x,\alpha)$ is convex and piecewise linear. The problem $\min_{(x,\alpha)} F_\beta^\mu(x,\alpha)$ in (4.12) can be approximated by the following:

$$\begin{aligned} \min_{x,\alpha} \quad & \alpha + \frac{1}{m(1-\beta)} \sum_{i=1}^{m} [-\mu_i^T x - \alpha]^+ \\ \text{s.t.} \quad & x \in \Omega \ . \end{aligned} \tag{4.17}$$

By introducing auxiliary real variables z_i for $i = 1, \ldots, m$, the problem (4.17) can be transformed into the following LP problem (Rockafellar and Uryasev [29]):

$$\begin{array}{lll} \min\limits_{x,z,\alpha} & \alpha + \dfrac{1}{m(1-\beta)} \displaystyle\sum_{i=1}^{m} z_i & \\ \text{s.t.} & x \in \Omega \ , & \\ & z_i \geq 0 \ , & \\ & z_i + \mu_i^T x + \alpha \geq 0, \ i = 1, \ldots, m \ . & \end{array} \tag{4.18}$$

Applying the approximation technique (4.18) for problem (4.14), our CVaR robust mean-variance portfolio optimization model (4.9) can be approximated by the following QP problem:

$$\begin{array}{lll} \min\limits_{x,z,\alpha} & \alpha + \dfrac{1}{m(1-\beta)} \displaystyle\sum_{i=1}^{m} z_i + \lambda x^T \bar{Q} x & \\ \text{s.t.} & x \in \Omega \ , & \\ & z_i \geq 0 \ , & \\ & z_i + \mu_i^T x + \alpha \geq 0, \ i = 1, \ldots, m \ , & \end{array} \tag{4.19}$$

where each μ_i is an independent sample of μ generated from an assumed distribution. This QP problem has $O(m+n)$ variables and $O(m+n)$ constraints, where m is the number of μ-samples and n is the number of assets.

Since the min-max robust MV model considers the worst sample scenario of μ in its uncertainty set S_μ, the choice of S_μ has significant impact on the optimal portfolio decision. This impact can be observed by comparing the performance of the min-max robust actual frontiers between Figure 3.2(a) and Figure 3.4(a). Unlike the min-max robust MV model, the CVaR robust MV model does not need any bounded uncertainty set for μ. Instead, it uses the probability bound β and optimizes a portfolio under the $(1-\beta)$ worst mean loss scenarios, which are obtained by generating a finite number of μ-samples from an assumed distribution. Therefore, the optimal portfolio generated by the CVaR robust MV model tends to be less sensitive to the possible outliers in the generated μ-samples. Note that, in spite of the common behaviors as compared to min-max robust portfolios, the CVaR robust portfolios generated under different μ distribution assumptions may perform differently. In Chapter 5, we will illustrate the performance differences when μ-samples are generated using two sampling techniques that are based on different μ distribution assumptions.

4.4.3 CVaR Robust MV Actual Frontiers

The CVaR robust MV model (4.9) presented in Section 4.4.1 considers minimizing the risk-adjusted CVaR_β^μ. In this formulation, both the CVaR^μ of mean loss and the return variance are included in the objective function, and λ is used as a risk-aversion parameter to adjust the trade-off between return risk (which is measured by variance) and CVaR_β^μ. When $\lambda = 0$, the resulting portfolio achieves the minimum CVaR_β^μ; when $\lambda = \infty$, the resulting portfolio achieves the minimum return risk. Note that, since the mean loss quantified by CVaR_β^μ is defined as the negative of the expected return, for all the CVaR robust portfolios generated under a particular confidence level β, the one with the minimum CVaR_β^μ tends to achieve the maximum (actual) expected return; following the same reason, the one with the minimum return risk tends to achieve the minimum (actual) expected return. Therefore, by specifying a confidence level β and solving problem (4.9) for all possible values of λ ranging from 0 to ∞, we obtain a CVaR robust MV actual frontier: it contains the entire set of CVaR robust portfolios under the confidence level β ranging from the maximum (actual) expected return to the minimum (actual) expected return.

4.4.4 Generating Mean Scenarios

With the CVaR minimization approach presented in Section 4.4.2, function $F_\beta^\mu(x, \alpha)$ is approximated by function $\bar{F}_\beta^\mu(x, \alpha)$, which considers the average of the $(1 - \beta)$-tail of the mean loss distribution. Generating mean loss scenarios depends on the specification of the μ distribution. However, this distribution is usually not known. In practice, one can use bootstrapping or resampling technique to generate some possible/reasonable realizations; the example in Section 3.3.1 utilizes the resampling technique. Here we consider another possible way to generate the μ distribution.

CHI Sampling Technique

Alternatively, we can generate samples that based on the statistics as described in (3.5), i.e., the quantity

$$\frac{T(T-n)}{(T-1)n}(\bar{\mu}-\mu)^T Q^{-1}(\bar{\mu}-\mu)$$

has a χ^2_n distribution with n degrees of freedom. see e.g., Garlappi et al. [12]. This sampling technique is subsequently referred to the CHI technique.

Given a random number $c \sim \chi^2_n$ associated with the quantity (3.5). The equation

$$\frac{T(T-n)}{(T-1)n}(\bar{\mu}-\mu)^T Q^{-1}(\bar{\mu}-\mu) = c \tag{4.20}$$

holds and can be transformed to

$$(\bar{\mu}-\mu)^T Q^{-1}(\bar{\mu}-\mu) = \phi \ , \tag{4.21}$$

where $\phi = \frac{(T-1)n}{T(T-n)}c$. Since Q is a symmetric positive definite matrix, using Cholesky factorization, it can be decomposed as

$$Q = GG^T, \tag{4.22}$$

where G is a lower triangular matrix with strictly positive diagonal entries, and G^T denotes the transpose of G. Substituting (4.22) into (4.21), we have:

$$(\bar{\mu} - \mu)^T (GG^T)^{-1} (\bar{\mu} - \mu) = \phi. \tag{4.23}$$

Applying matrix inversion and transformation properties on the left hand side, the equation in (4.23) can be re-arranged to:

$$(G^{-1}(\bar{\mu} - \mu))^T (G^{-1}(\bar{\mu} - \mu)) = \phi. \tag{4.24}$$

The left hand side of (4.24) specifies the square of the 2-norm of $(G^{-1}(\bar{\mu} - \mu))$. If we set

$$(G^{-1}(\bar{\mu} - \mu)) = y, \tag{4.25}$$

where y is a $n \times 1$ column vector, then (4.24) is equivalent to

$$\| y \|_2^2 = \phi. \tag{4.26}$$

The variable y in (4.26) can be considered as a point on the surface of a n-dimensional sphere whose radius equals to $\sqrt{\phi}$. Assuming that each y is uniformly distributed on the sphere surface, we can randomly choose a y using the normal-deviate method introduced in Muller [27] and Marsaglia [23]: Let $x = [x_1, x_2, \ldots, x_n]^T$ be the $n \times 1$ column vector, which contains n independent random numbers. Each random number is generated from the standard normal distribution. If $\| x \|_2^2 = s$, i.e.,

$$\sqrt{x_1^2 + x_2^2 + \ldots, + x_n^2} = \sqrt{s}, \tag{4.27}$$

then y can be obtained by:

$$y = \sqrt{\frac{\phi}{s}} x. \tag{4.28}$$

After getting y, we can compute the μ-sample by transforming the equation (4.25) to the form:

$$\mu = \bar{\mu} + Gy \tag{4.29}$$

If we generate m independent samples of c from the χ_n^2 distribution, then we obtain m independent samples of y. This gives m independent samples of μ. In Chapter 5, we investigate the characteristics of the CVaR robust actual frontier based on the CHI sampling technique, and demonstrate the statistical difference between the CHI and RS techniques.

4.5 Conclusion and Remarks

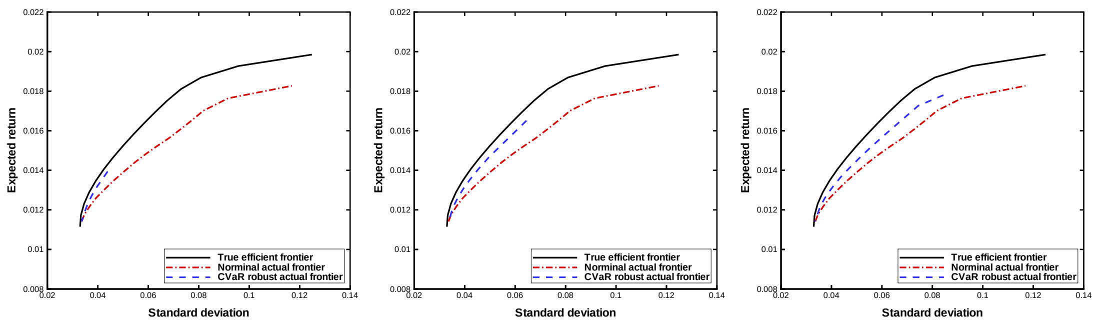
This chapter presents the CVaR robust MV portfolio optimization model. We show how this model adjusts the portfolio's conservatism level with respect to the estimation risk in mean return. In the following chapter, we conduct computational studies on the performance of CVaR robust portfolios. Comparisons are made with min-max robust portfolios in terms of actual frontier variation, portfolio efficiency, and portfolio diversification.

Chapter 5

Performance of CVaR Robust Portfolios

While the min-max robust MV model (3.3) is essentially quantile-based and focuses on the worst case scenario in an uncertainty set, the CVaR robust MV model (4.9) takes a distribution into consideration and ensures the best performance with respect to the average of the tail. We now compare min-max robust portfolios with CVaR robust portfolios in terms of actual frontier variation, portfolio efficiency, and portfolio diversification.

We first consider the 10-asset example used in Chapter 3. We generate 10,000 μ-samples using the RS (presented in Section 3.3.1) and CHI (presented in Section 4.4.4) technique as previously described. For each 10,000 sample set (which depends on the initial 100 return samples) of μ, we obtain a CVaR robust actual frontier by solving the CVaR robust MV model (4.19) for different values of $\lambda \geq 0$. For the 10-asset example using CHI-sampling, Figure 5.1(a)-(c) compares the CVaR robust actual frontiers (from the CVaR robust MV model) with the nominal actual frontiers (from the standard MV model based on the nominal estimates) for different confidence level β. We note that, unlike the min-max robust MV model (3.11) which uses the ellipsoidal uncertainty set (3.6), the CVaR robust actual frontiers mostly lie above the nominal actual frontiers.



(a) CHI: 90% confidence level (b) CHI: 60% confidence level (c) CHI: 30% confidence level

Figure 5.1: CVaR robust actual frontiers and nominal actual frontiers for the 10-asset example (in Table 3.1). CVaR robust actual frontiers are calculated based on 10,000 μ-samples generated via the CHI-sampling technique. Nominal actual frontiers are calculated by using the standard MV model with parameter $\bar{\mu}$ estimated based on 100 return samples.

To illustrate the characteristics of the CVaR robust actual frontier, we repeat the sampling procedure 100 times. For each set of 10,000 μ-samples, we compute three separate actual frontiers using $\beta = 90\%, 60\%$ and 30%. The top plots (a)–(c) in Figure 5.2 are for the RS technique, and the bottom plots (d)–(f) are for the CHI sampling technique. Note that the right-most point on a CVaR robust actual frontier corresponds to the maximum-return portfolio (with $\lambda = 0$). We describe the main observations in the following sections.

5.1 Sensitivity to Initial Data

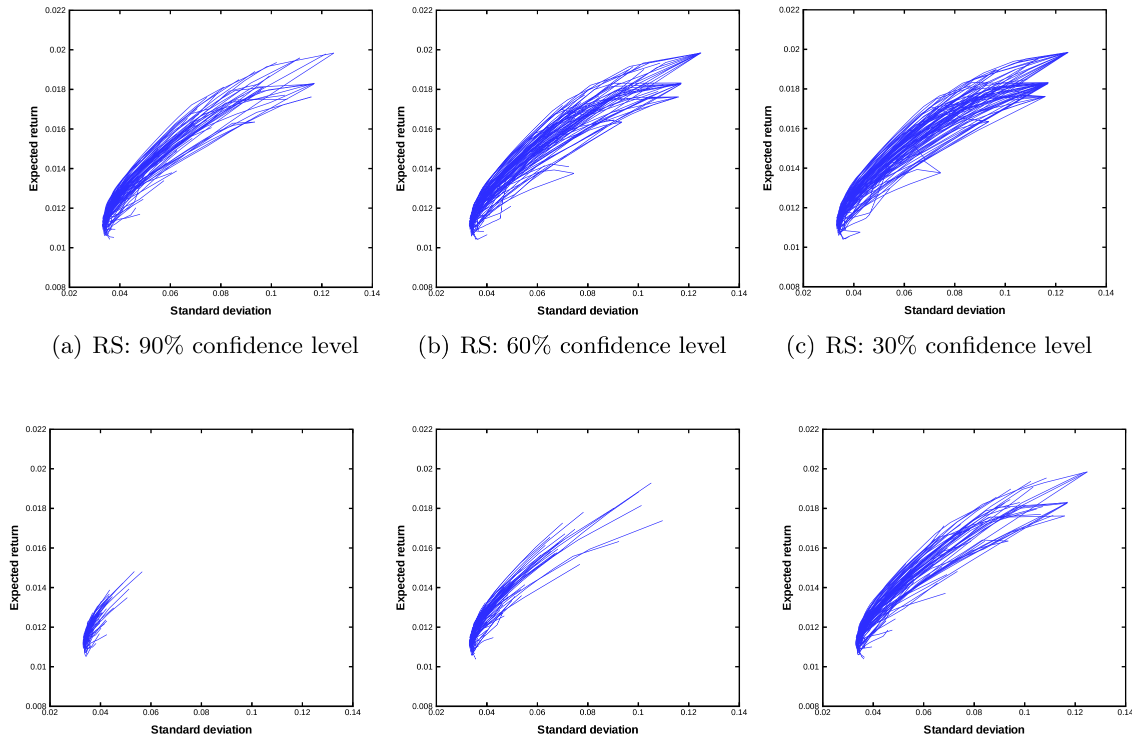
Similar to the min-max robust actual frontiers (with an interval uncertainty set) in Figure 3.4, the CVaR robust actual frontiers in Figure 5.2 also vary with the initial data used to generate each set of μ-samples. The variation of actual frontiers mainly comes from the variation in the estimate $\bar{\mu}$, which is computed based on 100 initial return samples. Since only a limited number of return samples are available in practice, variations inevitably exist in robust MV models, whether min-max robust or CVaR robust is considered.

The level of variation can be considered as an indicator of the level of estimation risk exposed by the portfolios from a robust model. It can be observed in Figure 5.2 that the variation seems to increase as the confidence level β decreases. This suggests that it may be reasonable to interpret β as an estimation risk aversion parameter: An investor who is more risk averse to estimation risk may choose a larger β. On the other hand, an investor who is more tolerant to estimation risk may choose a smaller β. The plots in Figure 5.2 depict the positive association between β and a portfolio's conservatism level.

In addition, we note that the variations of the actual frontiers in Figure 5.2(a)-(c) are larger than the ones in Figure 5.2(d)-(f). Figure B.1(a)-(h) in Appendix B compares the (marginal) distribution (10,000 mean return samples) for each of the 8 assets (from Table 2.1) generated using the RS and CHI sampling techniques. As can be seen, the samples obtained from the CHI technique have larger variances.

5.2 Adjustment of Portfolio's Conservative Level

In addition to variation in actual frontiers, we also evaluate the "average" performance of these CVaR robust actual frontiers. For Figure 5.2, we compute the "average" actual frontiers and plot them against the true efficient frontier in Figure 5.3. The true efficient frontier is used as a benchmark to assess the portfolio efficiency. The plots for the RS technique are on the top panel, while the ones for the CHI technique are on the bottom panel. As we can see, when β approaches 1, CVaR robust actual frontiers become shorter on "average" (i.e., the path length becomes smaller); the maximum expected return achievable becomes lower. As it is expected that an investor who is more averse to estimation risk obtains a smaller return, this confirms that it is reasonable to regard β as an indicator for the level of tolerance for estimation risk. On the other hand, an investor who is tolerant towards



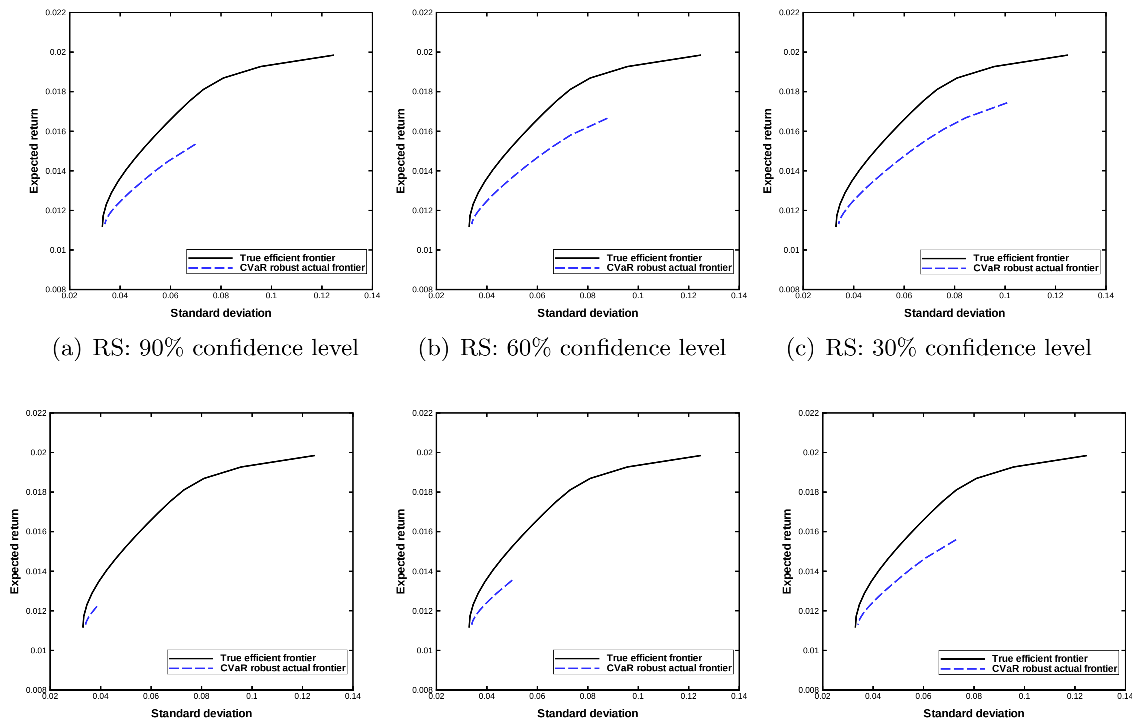
(a) RS: 90% confidence level (b) RS: 60% confidence level (c) RS: 30% confidence level

(d) CHI: 90% confidence level (e) CHI: 60% confidence level (f) CHI: 30% confidence level

Figure 5.2: 100 CVaR robust actual frontiers calculated based on 10,000 μ-samples. The true data is from Table 3.1.

estimation risk chooses a smaller β, the maximum expected return achievable becomes higher.

CVaR robust actual frontiers generated using the RS and the CHI sampling techniques seem to be different. For the same β value, the variations of the CVaR robust actual frontiers in Figure 5.2(d)–(f), corresponding to CHI, are dominated by the corresponding ones in Figure 5.2(a)–(c), corresponding to RS. This is likely due to different distributions generated by the two sampling techniques, see Appendix B. The CVaR robust actual frontiers for the two sampling techniques also have different performance in their "average" case. The "average" CVaR robust actual frontiers in Figure 5.3(d)–(f) achieve lower maximum expected returns than the corresponding ones in Figure 5.3(a)–(c). This happens because the μ-samples generated using the CHI technique have larger deviations, which leads to worse mean loss scenarios.

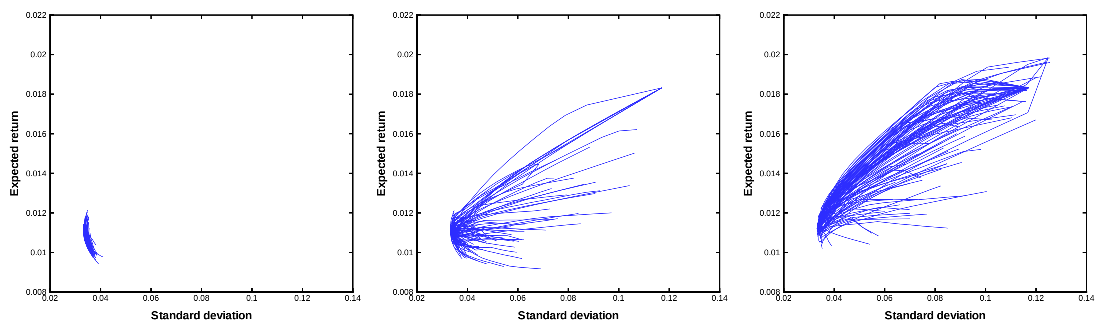


(a) RS: 90% confidence level (b) RS: 60% confidence level (c) RS: 30% confidence level

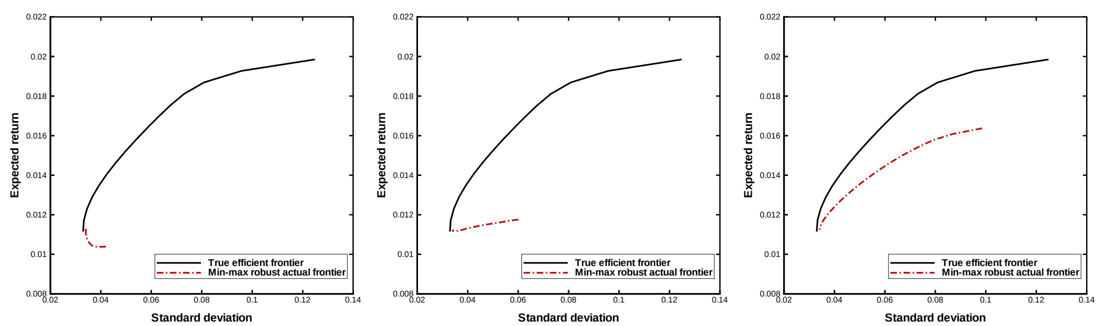
(d) CHI: 90% confidence level (e) CHI: 60% confidence level (f) CHI: 30% confidence level

Figure 5.3: "Average" CVaR robust actual frontiers calculated based on 10,000 μ-samples for the 10-asset example in Table 3.1.

It is also important to note that, although changing the confidence level β affects the maximum expected return achievable, the deviation of the CVaR robust actual frontiers from the true efficient frontier does not seem to be significantly affected. In addition, on "average", the deviation seems to be relatively insensitive for different sampling methods. On the other hand, the deviation (from the true efficient frontier) of the min-max robust actual frontiers varies significantly with the percentile value specifying μ^L. This can be observed from Figure 5.4(a)-(c) where the 100 min-max robust actual frontiers in each plot are computed based on different percentiles corresponding to μ^L. The μ-samples, based on which the percentiles are calculated, are generated using the CHI sampling technique. Note that here we use the same μ-samples as the ones used for generating the CVaR robust actual frontiers in Figure 5.2(d)-(f). As can be seen clearly, as the percentile value changes from 0 to 50, not only the variation but also the overall appearance of the min-max robust actual frontiers change significantly. This causes their "average" actual frontiers, which are plotted in Figure 5.4(d)-(f), have different deviations from the true efficient frontier. In addition, for this 10-asset example, the min-max robust actual frontiers in Figure 5.4(a)-(c) exhibit more variations in comparison with the CVaR robust actual frontiers in Figure 5.2(d)-(f).



(a) $\mu^L = 0$ percentile (b) $\mu^L = 25$ percentile (c) $\mu^L = 50$ percentile



(d) Average: 0 percentile (e) Average: 25 percentile (f) Average: 50 percentile

Figure 5.4: 100 min-max robust actual frontiers based on different percentiles for the 10-asset example in Table 3.1. μ-samples are generated using the CHI technique.

5.3 Portfolio Diversification

In this section, we demonstrate that under the estimation risk in μ, compared with the min-max robust portfolios, the CVaR robust portfolios are more diversified. We also numerically show that the diversification level of CVaR robust portfolios decreases as the confidence level β decreases.

5.3.1 Diversification Under Estimation Risk

An important way of minimizing the volatility in a portfolio is to diversify the portfolio. Portfolio diversification means spreading the total investment across a wide variety of asset classes, so the exposure to individual asset risk will be reduced. In statistical terms, while achieving the same expected return, one should diversify the portfolio so that a combined standard deviation of several assets is lower than the standard deviation of the individual asset.

Traditional MV Model

The traditional MV model (2.6) suggests how rational investors will use diversification to optimize their portfolios. As the risk-aversion parameter λ decreases, the level of diversification decreases. This will increase both the portfolio expected return and its associated return risk. When $\lambda = 0$, the portfolio typically achieves the maximum expected return by allocating all investment in the highest-return asset without considering the associated return risk. This portfolio (with $\lambda = 0$) is referred to as the maximum-return portfolio. In fact, even for $\lambda \neq 0$ but sufficiently small, the optimal MV portfolio tends to concentrate on a single asset. Given that the exact mean return is unknown, it means that the optimal MV portfolios can concentrate on a wrong asset due to estimation error. This can result in potentially disastrous performance in practice.

Min-max Robust MV Model

For the min-max robust MV model (3.3) with an interval uncertainty set for μ, the min-max robust portfolio is determined by the lower bound of the interval, μ^L, which is typically determined based on a confidence level. Thus, for the maximum-return portfolio computed from the min-max robust MV model, the allocation is still typically concentrated in a single asset. *Note that this is independent of the values of* μ^L. Moreover, due to estimation error, this allocation concentration may not necessarily result in a higher actual portfolio expected return. This is because that the asset which has the highest worst-case mean return in μ^L may not have the highest true mean return in μ. As an example, Figure 5.3 depicts that, on "average", the maximum expected return of the min-max robust actual frontier is significantly lower than that of the true efficient frontier.

CVaR Robust MV Model

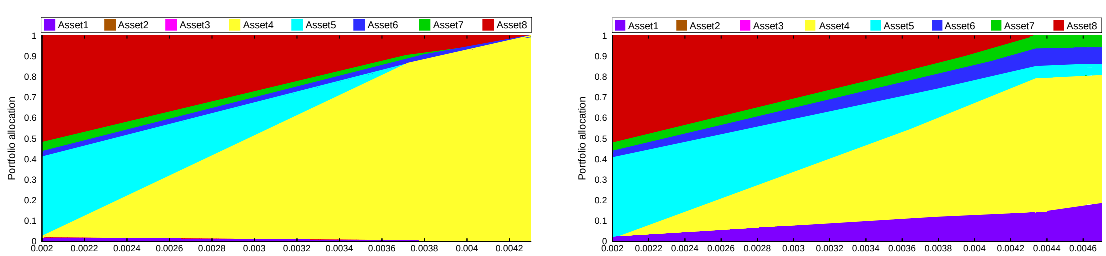
The CVaR robust MV model determines the optimal portfolio in a way that contrasts to the min-max robust MV model. Instead of focusing on the single worst case scenario μ^L of μ, the CVaR robust MV model optimizes a portfolio by considering the $(1-\beta)$-tail of the mean loss distribution. This forces the resulting portfolio to be more diversified. Therefore, even when ignoring return risk (i.e., $\lambda = 0$), the allocation of the CVaR robust portfolio (which typically achieves the maximum-return for the given β) is usually distributed among more than one asset, if β is not too small.

5.3.2 Computational Examples

Next we demonstrate that CVaR robust portfolios are more diversified, when compared with min-max robust portfolios. In addition, the diversification level decreases as the confidence level β decreases.

Composition of CVaR Robust Portfolios

Our first example illustrates the diversification property of the maximum-return portfolio (with $\lambda = 0$) computed from the CVaR robust MV model. We compute both the min-max robust and CVaR ($\beta = 90\%$) robust actual frontiers for the 8-asset example in Table 2.1. The computations are based on 10,000 μ-samples generated from the CHI sampling technique described in Section 4.4.4. Each frontier is formed by the portfolios computed using a sequence of λ ranging from 0 to 1000. Table 5.1(a) and 5.1(b) list the portfolio weights of the two actual frontiers for each λ value. When $\lambda = 0$, the maximum-return portfolio computed by the min-max robust MV model in Table 5.1(a) focuses all holdings in Asset4, whereas the one computed by the CVaR robust MV model in Table 5.1(b) are diversified into five different assets. We can also compare the composition graphs of the portfolios on the two actual frontiers. They are presented in Figure 5.5(a) and 5.5(b), respectively. For the minimum-return portfolio at the left-most end of each composition graph, most of the investment is allocated in Asset5 and Asset8. As the expected return value increases from left to right, both assets are gradually replaced by a mixture of other assets. However, close to the maximum-return end of the graphs, the compositions in Figure 5.5(b) are more diversified than that in Figure 5.5(a).



(a) Min-max robust portfolios

(b) CVaR robust ($\beta = 90\%$) portfolios

Figure 5.5: Compositions of min-max robust and CVaR robust ($\beta = 90\%$) portfolio weights for the 8-asset example in Table 2.1.

Table 5.1 Portfolio weights of min-max robust and CVaR robust ($\beta = 90\%$) actual frontiers for the 8-asset example in Table 2.1

(a) Min-max Robust Portfolios Weights

λ	Asset1	Asset2	Asset3	Asset4	Asset5	Asset6	Asset7	Asset8
0	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00
100	0.00	0.00	0.00	0.73	0.07	0.00	0.00	0.20
200	0.00	0.00	0.00	0.35	0.30	0.01	0.04	0.43
300	0.00	0.00	0.00	0.23	0.30	0.01	0.04	0.43
400	0.00	0.00	0.00	0.17	0.32	0.02	0.04	0.46
500	0.01	0.00	0.00	0.13	0.34	0.02	0.04	0.47
600	0.01	0.00	0.00	0.10	0.35	0.02	0.04	0.48
700	0.01	0.00	0.00	0.09	0.35	0.02	0.04	0.49
800	0.01	0.00	0.00	0.07	0.36	0.02	0.05	0.49
900	0.01	0.00	0.00	0.06	0.36	0.02	0.05	0.50
1000	0.01	0.00	0.00	0.05	0.37	0.02	0.05	0.50

(b) CVaR Robust ($\beta = 90\%$) Portfolios Weights

λ	Asset1	Asset2	Asset3	Asset4	Asset5	Asset6	Asset7	Asset8
0	0.18	0.00	0.00	0.63	0.05	0.08	0.06	0.00
100	0.05	0.00	0.00	0.16	0.30	0.04	0.06	0.39
200	0.04	0.00	0.00	0.09	0.34	0.04	0.06	0.44
300	0.04	0.00	0.00	0.06	0.36	0.03	0.05	0.47
400	0.03	0.00	0.00	0.04	0.37	0.03	0.05	0.48
500	0.03	0.00	0.00	0.03	0.37	0.03	0.05	0.49
600	0.03	0.00	0.00	0.02	0.37	0.03	0.05	0.49
700	0.03	0.00	0.00	0.02	0.38	0.03	0.05	0.50
800	0.03	0.00	0.00	0.01	0.38	0.03	0.05	0.50
900	0.03	0.00	0.00	0.01	0.38	0.03	0.05	0.50
1000	0.02	0.00	0.00	0.01	0.38	0.03	0.05	0.51

In Figure 5.7(a), the two actual frontiers corresponding to the composition graphs in Figure 5.5(a) and Figure 5.5(b) are plotted. The portfolios on the min-max robust actual frontier are less diversified and suffer larger return risk than that on the CVaR robust actual frontier for the same level of expected returns. Meanwhile, the maximum expected return of the min-max robust actual frontier is also lower. The min-max portfolio in this case allocates all investment in Asset4, whose true mean return is only 0.4734×10^{-2}. On the other hand, the CVaR robust actual frontier being higher and more to the left indicates the benefits of diversifying a portfolio under estimation risk.

Diversification

Next, we illustrate the impact of β on the level of diversification. As discussed in Chapter 4, in the CVaR robust MV model, β can represent an investor's estimation risk aversion level. The larger the β value, the more conservative the investor is with respect to estimation risk in μ. The CVaR robust actual frontier in Figure 5.7(a) is computed for $\beta = 90\%$. Now using the same dataset as in the first example, we compute the CVaR robust actual frontiers for $\beta = 60\%$ and $\beta = 30\%$, and list their portfolio weights in Table 5.2(a) and 5.2(b), respectively. The corresponding portfolios' composition graphs are also presented in Figure 5.6(a) and 5.6(b). By comparing the maximum-return portfolios ($\lambda = 0$) among Table 5.1(b), 5.2(a) and 5.2(b), the weights are less diversified as the value of β decreases. This effect can also be observed by comparing the compositions in Figure 5.5(b), 5.6(a) and 5.6(b). In particular, when $\lambda = 0$, the portfolio for $\beta = 30\%$ in Table 5.2(b) allocate all investment in a single asset. However, unlike the min-max robust portfolio in Table 5.1(a), which is concentrated on Asset4, this portfolio is concentrated in Asset1.

For the CVaR robust MV model, the relationship between the decrease in diversification and the decrease in β confirms that it is reasonable to regard β as a risk aversion parameter for estimation risk. An investor who is more risk averse to the estimation risk choose a

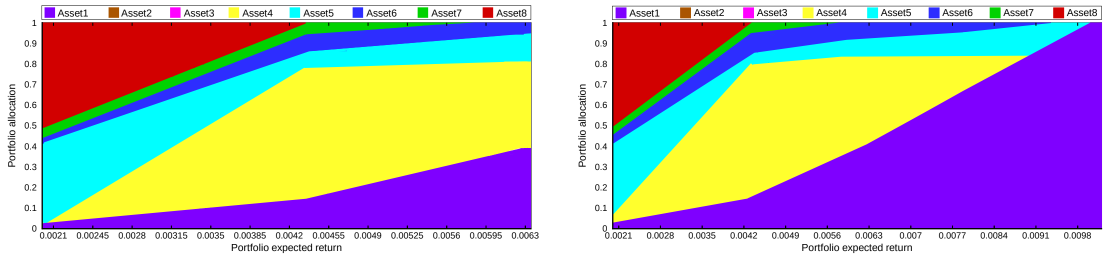
Table 5.2 Portfolio weights of CVaR robust (β =60%) and (β =30%) actual frontiers for the 8-asset example in Table 2.1

(a) CVaR robust (β =60%) Portfolios Weights

λ	Asset1	Asset2	Asset3	Asset4	Asset5	Asset6	Asset7	Asset8
0	0.39	0.00	0.00	0.42	0.13	0.06	0.00	0.00
100	0.06	0.00	0.00	0.21	0.28	0.05	0.06	0.35
200	0.04	0.00	0.00	0.11	0.33	0.04	0.06	0.43
300	0.04	0.00	0.00	0.07	0.35	0.03	0.05	0.46
400	0.03	0.00	0.00	0.05	0.36	0.03	0.05	0.47
500	0.03	0.00	0.00	0.04	0.37	0.03	0.05	0.48
600	0.03	0.00	0.00	0.03	0.37	0.03	0.05	0.49
700	0.03	0.00	0.00	0.02	0.38	0.03	0.05	0.49
800	0.03	0.00	0.00	0.02	0.38	0.03	0.05	0.50
900	0.03	0.00	0.00	0.01	0.38	0.03	0.05	0.50
1000	0.02	0.00	0.00	0.01	0.38	0.03	0.05	0.51

(b) CVaR robust (β =30%) Portfolios Weights

λ	Asset1	Asset2	Asset3	Asset4	Asset5	Asset6	Asset7	Asset8
0	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
100	0.07	0.00	0.00	0.25	0.26	0.05	0.06	0.31
200	0.05	0.00	0.00	0.12	0.32	0.04	0.05	0.41
300	0.04	0.00	0.00	0.08	0.35	0.03	0.05	0.45
400	0.03	0.00	0.00	0.05	0.36	0.03	0.05	0.47
500	0.03	0.00	0.00	0.04	0.37	0.03	0.05	0.48
600	0.03	0.00	0.00	0.03	0.37	0.03	0.05	0.49
700	0.03	0.00	0.00	0.02	0.38	0.03	0.05	0.49
800	0.03	0.00	0.00	0.02	0.38	0.03	0.05	0.50
900	0.03	0.00	0.00	0.01	0.38	0.03	0.05	0.50
1000	0.02	0.00	0.00	0.01	0.38	0.03	0.05	0.51

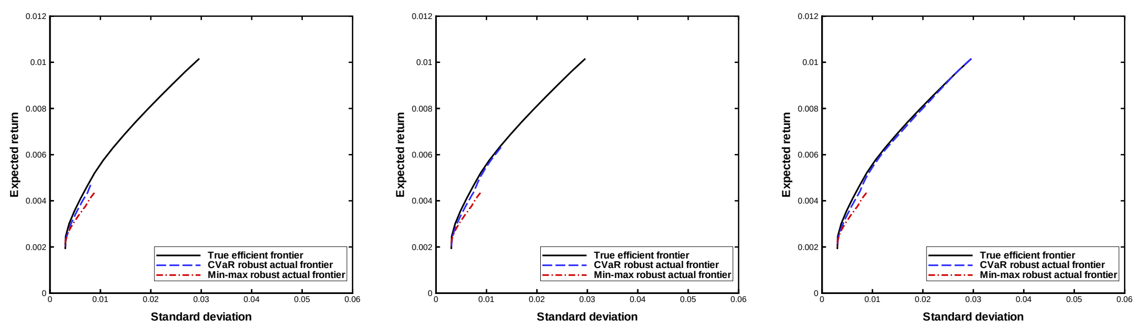


(a) CVaR robust (β =60%) portfolios

(b) CVaR robust (β =30%) portfolios

Figure 5.6: Compositions of CVaR robust (β =60%) and (β =30%) portfolio weights for the 8-asset example in Table 2.1.

larger β value and expect a more diversified portfolio. As discussed before, this portfolio may achieve a lower expected return. However, it also has less variations with respect to the initial data used to generate μ-samples. This reduces the possibility of the portfolio having disastrous performance when there exists a potentially large estimation risk of μ.



(a) 90% confidence level

(b) 60% confidence level

(c) 30% confidence level

Figure 5.7: Min-max robust and CVaR robust (β =90%, 60% and 30%) actual frontiers for the 8-asset example in Table 2.1

We can illustrate how the diversification level of a portfolio is affected by the tolerance level for estimation risk in μ in our CVaR robust MV model. We plot the actual frontiers for Table 5.2(a) and Table 5.2(b) in Figure 5.7(b) and 5.7(c), respectively. Compared with Figure 5.7(a) for $\beta = 90\%$, both the maximum expected return and the associated return risk increase as the β value decreases. Coincidentally, in this case, the maximum expected return in Figure 5.7(c) matches the exact solution obtained by using true parameter values on the traditional MV model (2.6). As demonstrated in Figure 5.2(f), there can be large variations on the compositions of the maximum-return portfolios for different runs, and hence obtaining the exact solution for $\lambda = 0$ would not happen for every case.

Table 5.3 Percentages of diversified maximum-return ($\lambda = 0$) portfolios

Confidence level β	0%	30%	60%	90%
Diversification Percentage (CHI)	0%	53%	85%	100%
Diversification Percentage (RS)	0%	18%	37%	64%

For the 100 CVaR robust actual frontiers in each of the six graphs in Figure 5.2, we compute the percentage of diversified maximum-return portfolios, and list them in Table 5.3. Here a portfolio is considered to be invested in an asset if its allocation percentage is greater than 1% and a portfolio is classified here as diversified if it consists of at least two assets. Comparing the percentages for different β values at each row, we conclude that the non-decreasing relationship between diversification level and β value generally holds regardless of the μ sampling techniques.

For the same β values, the diversification percentage of the RS technique is much smaller than that of the CHI technique. This may be due to the different distribution properties, as indicated in Figure B.1(a)-(h) in Appendix B.

5.4 Conclusion and Remarks

In this chapter, we analyze the performance characteristics of CVaR robust portfolios in terms of robustness, efficiency, and diversification. In the CVaR robust MV model, the confidence level β can be interpreted as an estimation risk aversion parameter. By changing the value of β, the conservatism of CVaR robust portfolios with respect to the estimation risk in mean return is adjusted. We also demonstrate through examples that the maximum-return portfolio (with $\lambda = 0$) from the CVaR robust MV model can be more diversified than the one from the min-max robust MV model, and the level of diversification decreases as the value of β decreases. Of course, simply allocating the investment into a large number of assets does not automatically make the portfolio diversified. The key is to ensure the assets have large varieties in their risk and return characteristics. Therefore, in practice, more detailed analysis on the nature of assets is needed for evaluating the diversification level of a portfolio.

Although the CVaR robust MV model is theoretically effective for alleviating portfolio estimation risk, the expensive computation may limit its applicability in practice. In next chapter, we introduce a more efficient method for computing CVaR robust portfolios.

Chapter 6

Efficient Technique for Computing CVaR Robust Portfolios

One potential disadvantage of the CVaR robust MV model (4.19) (which is a QP approach), in comparison to the min-max robust MV model (3.3), is that it may require more time to compute a CVaR robust portfolio than a min-max robust portfolio. This chapter addresses the computational issues for computing CVaR robust portfolios, and introduces a smoothing approach which is more efficient than the QP approach.

6.1 Quadratic Programming Approach

This chapter begins with a computational analysis for computing CVaR robust portfolios. In Chapter 4, we have shown that the CVaR robust MV model can be approximated by the QP problem (4.19):

$$\begin{array}{ll} \min\limits_{x,z,\alpha} & \alpha + \dfrac{1}{m(1-\beta)} \displaystyle\sum_{i=1}^{m} z_i + \lambda x^T \bar{Q} x \\ \text{s.t.} & x \in \Omega \ , \\ & z \geq 0 \ , \\ & z_i + \mu_i^T x + \alpha \geq 0, \ i = 1, \ldots, m \ , \end{array}$$

where each μ_i is an independent mean sample and m is the number of mean samples. Given a finite number of mean samples, this QP approach uses a piecewise linear function to approximate the continuous differentiable CVaR function. The more samples are used, the better approximation is achieved.

A convex QP is one of the simplest constrained optimization problem, and can be solved quickly using software such as MOSEK. However, the QP approach (4.19), similar to the LP approach (4.18) introduced by Rockafellar and Uryasev [29] for minimizing CVaR, can become very inefficient for large scale CVaR optimization problem. In this QP problem, generating a new sample adds an additional variable (and constraint). Therefore, for n risky assets and m samples, the problem has a total of $O(n+m)$ variables and $O(n+m)$ constraints. Alexander et al. [1] show that when the simplex method and the interior-point method are used in the LP approach, the computational cost can quickly become prohibitive as the number of samples and/or assets become large. In addition, the efficiency of MOSEK depends heavily on the sparsity structures of the QP problem. The QP problem in (4.19) has a large dense block whose size is determined by the number of samples and the number of assets; see Alexander et al. [1].

6.1.1 Computational Efficiency

We illustrate below the extent of the increase in computational cost when the scale of the QP problem becomes larger. These computational efficiency issues of the LP approach (4.18) for minimizing CVaR have been investigated in Alexander et al. [1]. However, the main difference is that the CVaR robust MV model (4.9) has an additional quadratic term $x^T Qx$ because variance is used as the return risk measure. In addition, the machine used in this study is different from the one used in Alexander et al. [1] and the computing platform and softwares are also different versions. The computation in this thesis is done in MATLAB version 7.3 for Windows XP, and run on a Pentium 4 CPU 3.00GHz machine with 1 GB RAM. QP problems are solved using the MOSEK Optimization Toolbox for MATLAB version 7.

An Example

We first consider computing the maximum-return portfolio, i.e., $\lambda = 0$; in this case the QP (4.19) becomes the LP (4.18). We compare the CPU time used by MOSEK to solve the LP for different asset examples with different sample sizes. The mean μ and the covariance matrix Q for the 8-asset example are taken from Table 2.1. The μ for the 50/148-asset example is obtained by generating 50/148 independent normally distributed random numbers and scaling the numbers into appropriate range. The Q for the 50/148-asset is obtained by constructing a random positive semi-definite matrix. We generate the μ-samples using both the RS technique and the CHI technique, and report the CPU time for both techniques. Table 6.1 illustrates the CPU time required for each combination of asset examples and sample sizes.

It can be clearly observed from Table 6.1 that, when using MOSEK, the computational cost increases quickly as the sample size and the number of assets increase. For examples, for each size of the RS samples, the CPU time required by the 50-asset example is at least

Table 6.1 CPU time for the QP approach when $\lambda = 0$: $\beta = 0.90$

	RS Tech (CPU sec)			CHI Tech (CPU sec)		
# samples	8 assets	50 assets	148 assets	8 assets	50 assets	148 assets
5000	0.41	1.84	9.77	0.39	1.75	7.06
10000	0.88	3.56	20.41	0.77	4.25	10.38
25000	2.78	9.17	32.69	2.56	10.83	34.97
50000	4.14	17.75	61.13	5.36	19.55	123.45

twice as the one required by the 8-asset example. When the size of the CHI samples is increased from 10,000 to 25,000, the CPU time is increased by at least 150% for each asset example.

6.1.2 Approximation Accuracy

The QP approach (4.19) uses a piecewise linear function to approximate the continuously differentiable CVaR function. When the number of μ-samples approaches infinity, the approximation approaches the exact value. However, as shown in Table 6.1, the associated computational cost increases significantly as the sample size increases. Therefore, given the trade-off between efficiency and accuracy, one has to decide the sample size using which the computation satisfies the requirement for speed and is within the tolerance for computational error.

Recall that when $\lambda = 0$, the QP approach (4.19) is reduced to the LP approach (4.18) which is used for solving the CVaR minimization problem. In the following experiment, we compare the computational error of the LP approach as the sample size is increased from 5000 to 25,000. The computational error is measured by the deviation of the CVaR values computed at different runs from their average value. Using each of the two sampling techniques, we generate the same number of μ-samples 100 times, and for each time, we calculate the minimum CVaR value based on the solution obtained from the LP approach. Let CVaR_{avg} be the average of these CVaR values, and CVaR_{std} be the standard deviation of these values from CVaR_{avg}. The computational error, CVaR_{err}, is given by

$$\text{CVaR}_{\text{err}} = \frac{\text{CVaR}_{\text{std}}}{\text{CVaR}_{\text{avg}}} \ . \tag{6.1}$$

Table 6.2 shows the CVaR_{err} values for different asset examples. Comparing the values among different sample sizes, we observe that CVaR_{err} decreases as the sample size increases. This means increasing the sample size improves the approximation accuracy. When the sample size is increased from 5000 to 10,000, there is a significant improvement on accuracy for most asset samples. For our examples, the CVaR_{err} (RS technique) of the 148-asset example is decreased by 92% (from 0.13% to 0.01%), and the CVaR_{err} (CHI technique) of the 8-asset example is decreased by 96% (from 13.04% to 0.52%). However, when the sample size becomes larger ($\geq 10,000$), the improvement on accuracy is much smaller. We can make comparisons with the increases in CPU time illustrated in Table 6.1. When the sample size is increased from 10,000 to 25,000, the largest improvement on accuracy is happened on the 50-asset example (RS technique), whose CVaR_{err} is decreased by 54% (from 0.44% to 0.20%); while the corresponding increase on the CPU time is 93.57% (from 9.17 sec to 17.75 sec). In fact we compare the CVaR_{err} change and the CPU time change for every asset example when the sample size is increased from 10,000 to 25,000, and observe that the improvement in accuracy is always smaller than the increase in CPU time. These examples indicate that one has to make a large sacrifice on efficiency for a small improvement on accuracy when calculating the CVaR using this QP approach.

Table 6.2 CVaR_{err} for the QP approach when $\lambda = 0$: $\beta = 0.90$

	RS Tech (%)			CHI Tech (%)		
# samples	8 assets	50 assets	148 assets	8 assets	50 assets	148 assets
5000	0.90	0.52	0.13	13.04	7.68	0.08
10000	0.47	0.44	0.01	0.52	1.94	0.01
25000	0.32	0.20	0.01	0.41	1.70	0.01

The above experimental results demonstrate how the computational cost for solving the QP problem (4.19) increases as the sample size and the number of assets are increased. Compared to accuracy, computational cost increases much faster when the sample size becomes larger. This indicates that using the QP approach to compute CVaR robust portfolios can be very inefficient in practice.

6.2 Smoothing Approach

As an alternative to the QP approach (4.19) discussed in Section 6.1, we can compute CVaR robust portfolios more efficiently via the smoothing technique proposed by Alexander et al. [1]. It has been shown in Alexander et al. [1] that the smoothing technique directly exploits the structure of the CVaR minimization problem, and is computationally more efficient than the LP method. Next we investigate the computational performance comparison between the QP approach and the smoothing approach for computing CVaR robust portfolios.

As mentioned in Section 4.4.2,

$$\min_{x} \text{CVaR}_{\beta}^{\mu}(x) + \lambda x^T Q x \equiv \min_{x,\alpha} F_{\beta}^{\mu}(x,\alpha) + \lambda x^T Q x \ ,$$

where the function $F_{\beta}^{\mu}(x,\alpha)$, which is defined as

$$F_{\beta}^{\mu}(x,\alpha) = \alpha + \frac{1}{1-\beta} \int_{\mu \in R^n} [f(x,\mu) - \alpha]^+ p(\mu) \, d\mu \ , \tag{6.2}$$

is both convex and continuously differentiable.

The QP approach (4.19) approximates the function $F_{\beta}^{\mu}(x,\alpha)$ by the following piecewise linear objective function:

$$\bar{F}_{\beta}^{\mu}(x,\alpha) = \alpha + \frac{1}{m(1-\beta)} \sum_{i=1}^{m} [-\mu_i^T x - \alpha]^+ \ , \tag{6.3}$$

where each μ_i is a mean vector sample. When the number of μ-samples increases to infinity, the approximation approaches to the exact function.

Instead of using $\bar{F}^{\mu}_{\beta}(x, \alpha)$, Alexander et al. [1] suggest a piecewise quadratic function $\tilde{F}^{\mu}_{\beta}(x, \alpha)$ to approximate $F^{\mu}_{\beta}(x, \alpha)$. Let

$$\tilde{F}^{\mu}_{\beta}(x, \alpha) = \alpha + \frac{1}{m(1-\beta)} \sum_{i=1}^{m} \rho_{\epsilon}(-\mu_i^T x - \alpha), \tag{6.4}$$

where $\rho_{\epsilon}(z)$ is defined as:

$$\begin{cases} z & \text{if } z \geq \epsilon \\ \frac{z^2}{4\epsilon} + \frac{1}{2}z + \frac{1}{4}\epsilon & \text{if } -\epsilon \leq z \leq \epsilon \\ 0 & \text{otherwise.} \end{cases} \tag{6.5}$$

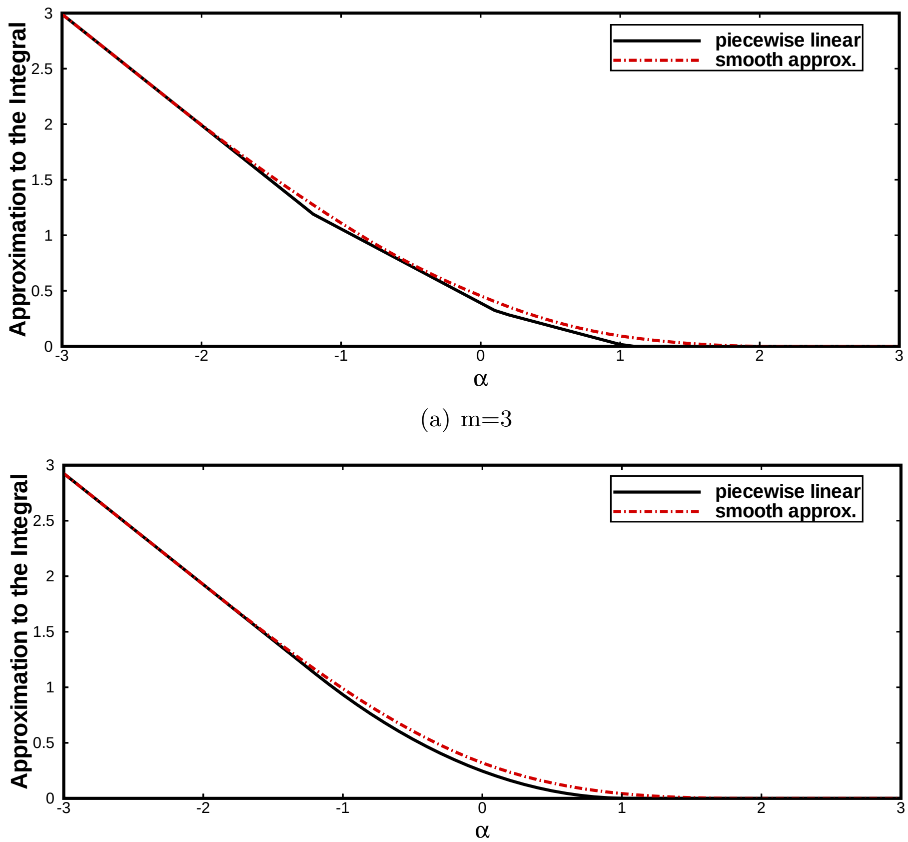
For a given resolution parameter $\epsilon > 0$, $\rho_{\epsilon}(z)$ is continuous differentiable, and approximates the piecewise linear function $\max(z, 0)$.

Figure 6.1 graphically compares the integral approximation achieved by functions $\frac{1}{m}\sum_{i=1}^{m}[S_i - \alpha]^+$ and $\frac{1}{m}\sum_{i=1}^{m}\rho_{\epsilon}(S_i - \alpha)$, where α and S_i are chosen from the range $[-3, 3]$ and $[-1.2, 1.05]$ respectively; Figure 6.1(a) is for $m = 3$ and Figure 6.1(b) is for $m = 10,000$. As can be observed, as the number of independent samples m increases, the difference between the two functions becomes smaller.

Applying the smoothing formulation (6.4), the CVaR robust MV model (4.9) can be formulated as the following problem:

$$\begin{array}{ll} \min\limits_{x,\alpha} & \alpha + \frac{1}{m(1-\beta)} \sum_{i=1}^{m} \rho_{\epsilon}(-\mu_i^T x - \alpha) + \lambda x^T \bar{Q} x \\ \text{s.t.} & x \in \Omega \ , \end{array} \tag{6.6}$$

where $\lambda \geq 0$ is a risk-aversion parameter. Note that when $\lambda = 0$, investors want to minimize the $(1-\beta)$-tail of the portfolio mean loss distribution without considering the associated



(a) m=3

(b) m=10,000

Figure 6.1: Approximation comparison between piecewise linear function $\frac{1}{m}\sum_{i=1}^{m}[S_i-\alpha]^+$ and smooth function $\frac{1}{m}\sum_{i=1}^{m}\rho_\epsilon(S_i-\alpha)$ with $\epsilon=1$.

risk. In this case, the term $\lambda x^T \bar{Q} x$ is eliminated, and the problem (6.6) is reduced to the CVaR minimization problem via the smoothing technique.

While QP formulation (4.19) has a total of $O(n+m)$ variables and $O(n+m)$ constraints, the smoothing formulation (6.6) only has $O(n)$ variables and $O(n)$ constraints. Therefore, increasing the sample size m does not change the number of variables (and constraints). We illustrate in the following that the smoothing approach significantly reduces the computational cost required for computing CVaR robust portfolios.

6.3 Comparisons Between the QP and Smoothing Approaches

6.3.1 Computational Efficiency

In Table 6.3, we report the CPU time required by the smoothing approach (6.6) for the same example in Table 6.1, which is included again for the ease of comparison. The smoothing approach is implemented based on the interior-point method introduced by Coleman and Li [10] for nonlinear minimization with bound constraints. The computation is done for both the RS and CHI sampling techniques, for which the CPU time is illustrated in Table 6.3(a) and 6.3(b), respectively.

The Case $\lambda = 0$

Comparing the CPU time between the two approaches, we observe that the smoothing approach is much more efficient than the QP approach for both sampling techniques. The problem of 148 assets and 25,000 samples can now be solved in less than 11 CPU seconds via the smoothing approach; while the same problem is solved in more than 30 CPU seconds via the QP approach. The CPU efficiency gap increases as the scale of the optimization problem (including the sample size and the number of assets) becomes larger. For 8 assets and 5000 samples, there is a small difference between the CPU time used by the two approaches. However, when the number of assets exceeds 50 and the sample size exceeds 5000, the difference becomes significant. All of these comparisons show that the smoothing approach achieves better computational efficiency.

Table 6.3 CPU time for computing maximum-return portfolios ($\lambda = 0$) MOSEK vs. Smoothing ($\epsilon = 0.005$): $\beta = 0.90$

(a) RS Technique

	MOSEK (CPU sec)			Smoothing (CPU sec)		
# samples	8 assets	50 assets	148 assets	8 assets	50 assets	148 assets
5000	0.41	1.84	9.77	0.34	0.50	2.55
10000	0.88	3.56	20.41	0.56	1.34	4.08
25000	2.78	9.17	32.69	1.22	3.28	8.11
50000	4.14	17.75	61.13	2.34	6.77	20.05

(b) CHI Technique

	MOSEK (CPU sec)			Smoothing (CPU sec)		
# samples	8 assets	50 assets	148 assets	8 assets	50 assets	148 assets
5000	0.39	1.75	7.06	0.42	0.34	1.98
10000	0.77	4.25	10.38	0.75	0.50	4.13
25000	2.56	10.83	34.97	1.77	1.36	10.25
50000	5.36	19.55	123.45	1.81	3.61	18.52

The Case $\lambda \geq 0$

Next we compare the CPU time between the QP approach and the smoothing approach for different λ values used in the CVaR robust MV model. Here we are interested in the CPU time difference not only between the two approaches, but also among different λ values of the same approach. Using four different λ values, Table 6.4 illustrate the CPU time required by both approaches. The 148-asset example is used for this experiment and μ-samples are generated using the RS technique. It can be observed that, with the same sample size, the smoothing approach is more efficient for each λ value. In addition, the CPU time required by the QP approach varies significantly with different λ values; while the one required by the smoothing approach is relatively insensitive to the value of λ. The best CPU efficiency of the QP approach is obtained for $\lambda = 0$ (when the covariance matrix Q is not used for optimization). However, when λ values are positive, the CPU time is significantly increased.

For example, the CPU time required by the QP approach for $\lambda = 0.1$ is about twice as much as the one for $\lambda = 0$ when 10,000 samples are used, and is about three times when 25,000 samples are used. This clearly indicates the increase in computational cost when Q is involved in the computation. On the other hand, this impact is minor for the smoothing approach, and sometimes the CPU time for positive λ values is even less than the one for $\lambda = 0$. Note that the increase in positive λ values does not necessary result an increase in CPU time. As an example, when 10,000 samples are used, both approaches require less CPU time for $\lambda = 1000$ than that for $\lambda = 0.1$.

Table 6.4 CPU time for different λ values ($\epsilon = 0.005$) for the 148-asset example: $\beta = 0.90$

	MOSEK (CPU sec)				Smoothing (CPU sec)			
# samples	0	0.1	10	1000	0	0.1	10	1000
5000	10.42	11.13	14.75	15.19	2.31	2.16	2.14	2.58
10000	18.33	42.77	29.41	36.66	3.70	3.55	4.00	3.36
25000	29.59	89.06	95.31	122.72	7.66	7.95	7.16	7.58
50000	56.36	163.17	202.17	210.78	21.28	21.55	19.27	17.41

6.3.2 Approximation Accuracy

In addition to computational efficiency, Alexander et al. [1] analyze the accuracy of the smoothing technique when it is applied for minimizing the CVaR for a portfolio of derivatives. They compare the difference between the CVaR values computed based on the LP approach implemented with MOSEK and the smoothing approach, and examine the impact on this difference when the value of the resolution parameter ε in (6.6) is changed. Here we conduct the same analysis for our CVaR robust MV model (6.6) that applies the smoothing technique. We determine the following relative difference in the CVaR^{μ} value computed via the smoothing approach:

$$Q_{\text{CVaR}^{\mu}} = \frac{\text{CVaR}_s^{\mu} - \text{CVaR}_m^{\mu}}{\mid \text{CVaR}_m^{\mu} \mid}\ , \tag{6.7}$$

where CVaR_s^μ and CVaR_m^μ are the CVaR^μ values obtained by using the QP approach and the smoothing approach, respectively. For comparison purpose, each pair of CVaR_s^μ and CVaR_m^μ are computed based on the same set of μ-samples generated from the RS technique. In addition, the confidence level for CVaR^μ, β, remains fixed and is set to 0.90.

For our analysis, we compare the Q_{CVaR^μ} for different sample sizes (including 10,000, 25,000 and 50,000), and determine the change on its value when sample size is increased. The effect of using the smoothing approach for approximation depends on the resolution parameter ϵ. Alexander et al. [1] suggest that the value of ϵ should be chosen between 0.05 to 0.005, and ϵ should be smaller for a larger sample size since this leads to a better approximation. The ϵ parameters used in our analysis are 0.005, 0.001 and 0.0005, and the CPU efficiency for each ϵ is evaluated.

Table 6.5 compares both Q_{CVaR^μ} and CPU time for different sample sizes and ϵ values. As expected, given the same ϵ, the absolute value of Q_{CVaR^μ} decreases when the sample size increases. This indicates that the difference between the CVaR^μ values approximated by the two approaches become smaller. The difference for the same sample size also becomes smaller when ϵ is decreased from 0.005 to 0.001, however, tends to be unchanged when ϵ is decreased from 0.001 to 0.0005. This shows that when the value of ϵ is small, the sample size has to increase much faster to make a noticeable decrease in the approximation difference. Note that when $\epsilon = 0$, the problem (6.6) with a finite number of samples is no longer smooth, i.e., the objective function $\tilde{F}_\beta^\mu(x, \alpha)$ is no longer continuously differentiable. Therefore, in the smoothing approach, ϵ can never be set to 0 and demonstrating the convergence on Q_{CVaR^μ} for ϵ approaches 0 will be computationally difficult.

It is expected that increasing the sample size causes the increase in CPU time. This is the expense for obtaining a better approximation. However, according to the comparisons made on CPU time in Table 6.3, this computational cost has been significantly decreased by the smoothing approach.

Table 6.5 Comparison of the CVaR^{μ} values computed by MOSEK and the proposed smoothing technique for different resolution parameter ϵ, $\beta = 95\%$ and $\lambda = 0$

(a) $\epsilon = 0.005$

	50 assets		148 assets		200 assets	
# samples	$Q_{\text{CVaR}}(\%)$	CPU sec	$Q_{\text{CVaR}}(\%)$	CPU sec	$Q_{\text{CVaR}}(\%)$	CPU sec
10000	-1.1225	2.61	-0.2253	3.83	-0.2260	8.81
25000	-0.0939	5.09	-0.0889	10.58	-0.0883	23.41
50000	-0.0513	5.11	-0.0459	19.39	-0.0472	44.86

(b) $\epsilon = 0.001$

	50 assets		148 assets		200 assets	
# samples	$Q_{\text{CVaR}}(\%)$	CPU sec	$Q_{\text{CVaR}}(\%)$	CPU sec	$Q_{\text{CVaR}}(\%)$	CPU sec
10000	-0.2974	1.86	-0.2236	2.44	-0.2234	4.28
25000	-0.0934	3.00	-0.0882	5.59	-0.0880	10.44
50000	-0.0504	4.14	-0.0454	12.52	-0.0466	29.20

(c) $\epsilon = 0.0005$

	50 assets		148 assets		200 assets	
# samples	$Q_{\text{CVaR}}(\%)$	CPU sec	$Q_{\text{CVaR}}(\%)$	CPU sec	$Q_{\text{CVaR}}(\%)$	CPU sec
10000	-0.2784	3.50	-0.2236	2.91	-0.2231	3.58
25000	-0.0934	3.80	-0.0882	5.19	-0.0879	8.16
50000	-0.0504	4.67	-0.0454	13.22	-0.0466	16.88

6.4 Conclusion and Remarks

In this chapter, a smoothing technique is implemented for computing CVaR robust portfolios. Unlike the QP approach, which uses a piecewise linear function to approximate the CVaR function, the smoothing technique uses a piecewise quadratic function which is continuous differentiable. As the number of μ-samples increases, the smoothing approximation approaches the CVaR function. Comparisons on computational efficiency and approximation accuracy are made between the two approaches. We show that, the smoothing approach is more computationally efficient in terms of CPU time for computing CVaR robust portfolios. In addition, when choosing appropriate resolution parameters and sample sizes, the CVaR^{μ} values obtained by the smoothing approach can be very close to the one obtained by the QP approach.

Chapter 7

Conclusion and Future Work

7.1 Conclusion

The classical mean-variance (MV) portfolio optimization model is typically based on the nominal estimates of mean returns and a covariance matrix from a set of return samples. Given that the number of return samples is limited in practice, the resulting optimal portfolios can vary significantly with the set of initial return samples; the actual performance of the MV efficient frontier can be potentially very poor. In this thesis, we investigate the estimation risk in the MV model and how it is addressed in a robust MV model. We consider estimation risk only in the mean returns and ignore that in the covariance matrix.

Recently, min-max robust mean-variance portfolio optimization has been proposed to address the estimation risk. With an ellipsoidal uncertainty set based on the statistics of the sample mean estimates, the resulting portfolio from the min-max robust MV model equals to the one from the standard MV model based on the nominal mean estimate but with a larger risk aversion parameter. We show that, with an interval uncertainty set $[\mu^L, \mu^U]$, the resulting min-max robust portfolio is essentially the MV optimal portfolio generated based on the lower bound μ^L. Of course, the min-max robust optimization problem becomes more complex when other types of uncertainty sets are used. But the min-max robust MV model, by nature, emphasizes the best performance under the worst-case scenario. In addition, it is difficult to select the appropriate uncertainty set in general. The min-max robust portfolio also ignores any probability information in the uncertain data.

We show that the min-max robust portfolio can also be very sensitive to the initial data used to generate an uncertainty set. In addition, if μ^L corresponds to the worst possible scenario, the min-max robust portfolio can be conservative and unable to achieve a sufficiently high expected return. Adjustment of the level of conservatism in the min-max robust MV model can be achieved by excluding bad scenarios from the uncertainty sets; but this is philosophically unappealing.

Given the existence of estimation risk, certain level of variation in actual frontiers (even from robust methods) are inevitable. However, due to smaller estimation error in the covariance matrix, variance-based actual frontiers tend to have small variations over the entire frontiers. Furthermore, we show, via examples, that the variance-based actual frontiers can sometimes be more efficient than the min-max robust actual frontiers. Thus the deviation of actual frontiers from the true (unknown in practice) efficient frontier is also important. In addition, proper mechanism in adjusting the level of conservatism is crucial in practice.

We propose a CVaR robust MV portfolio optimization model to address the estimation risk in mean return. In this model, a robust portfolio is determined based on a tail of worse portfolio mean loss scenarios, rather than nominal estimates (as in classical MV) or a single worst-case scenario (as in min-max robust). When the confidence level β is high, CVaR robust optimization focuses on a small set of extreme mean loss scenarios. The resulting portfolios are optimal against the average of these extreme mean loss scenarios and tend to be more conservative with respect to estimation risk.

More aggressive robust portfolios can be generated with a smaller confidence level β in the CVaR robust MV model. In contrast to the min-max robust MV model, the decrease in the level of conservatism is achieved by including a larger set of better mean loss scenarios; this results in less focus on the extreme poor scenarios. Decreasing the confidence level β corresponds to more acceptance of estimation risk, and our computational results suggest that there is little variation in the efficiency of the CVaR robust actual frontiers. Indeed, it seems reasonable to regard β as a risk aversion parameter for estimation risk.

When the uncertainty set is determined based on a quantile of the uncertain parameters with respect to an assumed distribution, the min-max robust MV model is essentially quantile-based. The CVaR robust MV model, on the other hand, is tail-based. Because of this, there are some crucial differences in the diversification of the robust portfolios generated from the two models. For example, in spite of the robust objective, the investment allocation for the min-max robust portfolio with $\lambda = 0$ (which achieves the maximum expected return) typically concentrates on a single asset, no matter what quantile value is chosen to be μ^L. The corresponding CVaR robust portfolio, on the other hand, typically consists of multiple assets for a high confidence level, e.g., $\beta = 90\%$. The level of diversification decreases as the value of β decreases. When $\beta = 0$, the CVaR robust portfolio with $\lambda = 0$ typically consists of a single asset as well.

Both the min-max robust and CVaR robust MV models are based on the distribution information of mean returns. However, this information may not be known precisely in practice. There are however statistical results and heuristic sampling techniques to generate some distributions for the uncertain parameters. In this thesis, we consider a RS-sampling technique and a CHI-sampling technique based on statistics of the parameter estimates. We demonstrate through computational examples that, using the two different sampling techniques, the characteristics of the CVaR robust actual frontiers obtained are similar.

Finally, we investigate the computational issue of the CVaR robust MV model, and implement a smoothing technique for computing CVaR robust portfolios. Unlike the QP approach, which uses a piecewise linear function to approximate the CVaR function, the smoothing approach uses a piecewise quadratic function that is continuously differential. We show that the smoothing approach is computationally more efficient for computing CVaR robust portfolios. In addition, as the number of mean return samples increases, the difference between the CVaR values approximated by the two approaches become smaller.

7.2 Possible Future Work

There are several potential extensions to our work. First, compared with the covariance matrix Q, the estimation error in the mean return μ is typically much larger, and the impact on the optimal portfolio selection is more severe; see Broadie [7] and Merton [24]. For this reason, we address the estimation risk in μ only and ignore that in Q. One possible future work is to extend the CVaR robust MV model to address the estimation risk in Q as well, and investigate possible techniques to compute the resulting CVaR robust portfolios.

Second, computing CVaR robust portfolio requires the distribution information of μ. We generate the distribution of μ by utilizing two sampling techniques that are based on different μ distribution assumptions: one is the RS technique and the other is the CHI technique. It would also be interesting to investigate the characteristics of CVaR robust portfolios that are based on other sampling techniques (such as bootstrapping and moving average). In particular, it is necessary to verify the performance of CVaR robust actual frontiers (such as: robustness, portfolio efficiency and portfolio diversification) are consistent for different μ sampling techniques.

Finally, we have shown in Chapter 6 that the smoothing approach is more efficient than the QP approach for computing CVaR robust portfolios. However, due to the scenario-based nature, the smoothing approach can still be time-consuming, especially when compared with the min-max robust MV model. Therefore, there is still room for enhancing the computational performance of the smoothing approach. For example, the current starting point of the interior-point method implemented by the smoothing technique is the middle of the convex feasible region. Since the choice of the starting point can have significant impact on the speed of convergence, effective algorithms can be explored to determine the best starting point so that the computation efficiency of the smoothing approach can be improved.

Appendix A

Theorems and Proofs

This section provides Theorem A.1 (and its proof) in Zhu et al. [33] that is used for proving Theorem 3.1.

Theorem A.1. *Assume that* $\bar{Q}$ *is symmetric positive definite and* $\chi \geq 0$. *The min-max robust portfolio for problem*

$$\begin{aligned} \min_{x} \quad & \max_{\mu} -\mu^T x + \lambda\sqrt{x^T \bar{Q} x} \\ s.t. \quad & (\bar{\mu} - \mu)^T \bar{Q}^{-1} (\bar{\mu} - \mu) \leq \chi \\ & e^T x = 1, \quad x \geq 0 \end{aligned} \tag{A.1}$$

is an optimal portfolio of the mean-standard deviation problem

$$\begin{aligned} \min_{x} \quad & -\mu^T x + \lambda\sqrt{x^T Q x} \\ s.t. \quad & e^T x = 1, \quad x \geq 0 \end{aligned} \tag{A.2}$$

with nominal estimates $\bar{\mu}$ *and* $\bar{Q}$ *for a larger risk aversion parameter* $\lambda + \sqrt{\chi}$.

Proof. Firstly we note that $x = 0$ is not a feasible point for (3.11).

For any $x \neq 0$, let μ^* be the minimizer of the inner optimization problem in (3.11) with respect to μ, i.e., μ^* solves

$$\begin{aligned} &\min_{\mu} \mu^T x \\ \text{s.t.} \quad &(\bar{\mu} - \mu)^T \bar{Q}^{-1} (\bar{\mu} - \mu) \leq \chi \; . \end{aligned}$$

Then there exists some $\rho < 0$ such that

$$x - \rho \bar{Q}^{-1} (\mu^* - \bar{\mu}) = 0 \; .$$

Thus

$$\mu^* = \bar{\rho} \bar{Q} x + \bar{\mu}, \qquad \text{where } \bar{\rho} = \frac{1}{\rho} < 0.$$

From

$$\bar{Q}^{-\frac{1}{2}} (\mu^* - \bar{\mu}) = \bar{\rho} \bar{Q}^{\frac{1}{2}} x$$

and

$$(\bar{\mu} - \mu^*)^T \bar{Q}^{-1} (\bar{\mu} - \mu^*) = \chi \; ,$$

we have

$$\bar{\rho}^2 = \frac{\chi}{x^T \bar{Q} x} \quad \text{and} \quad \bar{\rho} = -\frac{\sqrt{\chi}}{\sqrt{x^T \bar{Q} x}} \; .$$

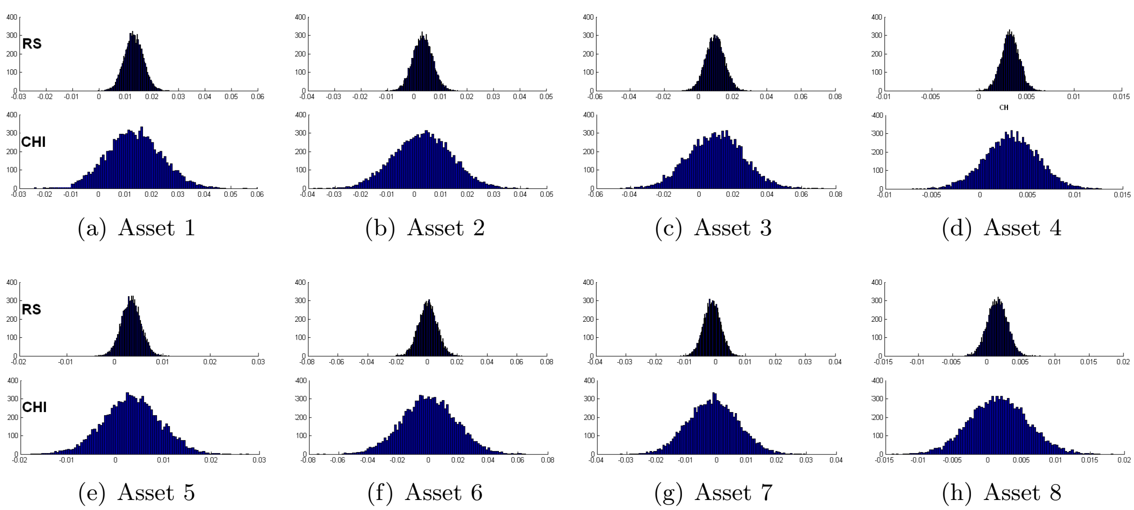
Thus the min-max robust mean-standard deviation portfolio can be obtained from

$$\begin{aligned} &\min_{x} -\bar{\mu}^T x + (\lambda + \sqrt{\chi}) \sqrt{x^T \bar{Q} x} \\ \text{s.t.} \quad &e^T x = 1, \quad x \geq 0 \; . \end{aligned}$$

This completes the proof. □

Appendix B

Distributions from RS and CHI Sampling Technique



(a) Asset 1 (b) Asset 2 (c) Asset 3 (d) Asset 4

(e) Asset 5 (f) Asset 6 (g) Asset 7 (h) Asset 8

Figure B.1: Distribution of mean return samples generated by sampling techniques RS(top) and CHI(bottom) for each asset in Table 2.1.

Bibliography

[1] S. Alexander, T. F. Coleman, and Yuying Li. Minimizing var and cvar for a portfolio of derivatives. *Journal of Banking and Finance*, 30(2):583–605, 2006.

[2] P. Artzner, F. Delbaen, J. Eber, and D. Heath. Thinking coherently. *Risk*, 10:68–71, 1997.

[3] A. Ben-Tal and A. S. Nemirovski. Robust convex optimization. *Math. of Oper. Res.*, 23(4):769805, 1998.

[4] M.J. Best. An algorithm for the solution of the parametric quadratic programming problem. *Applied Mathematics and Parallel Computing - Festschrift for Klaus Ritter, H. Fischer, B. Riedmüller and S. Schäffler (editors), Heidelburg: Physica-Verlag*, pages 57–76, 1996.

[5] M.J. Best and R.R. Grauer. On the sensitivity of mean-variance-efficient portfolios to changes in asset means: some analytical and computational results. *The Review of Financial Studies*, 4:315–342, 1991.

[6] F. Black and R. Litterman. Global portfolio optimization. *Financial Analysts Journal*, 48:28–43, 1992.

[7] M. Broadie. Computing efficient frontiers using estimated parameters. *Annals of Operations Research*, 45:21–58, 1993.

[8] F. Cavadini, A. Sbuelz, and F. Trojani. A simplified way of incorporating model risk, estimation risk and robustness in mean variance portfolio management. Working Paper, Tilburg University, 2001.

[9] V. K. Chopra and W. T. Ziemba. The effect of errors in means, variances, and co-variances on optimal portfolio choice. *Journal of Portfolio Management*, 19(2):6–11, 1993.

[10] T. F. Coleman and Yuying Li. An interior, trust region approach for nonlinear minimization subject to bounds. *SIAM Journal on Optimization*, 6(2):418–445, 1996.

[11] V. DeMiguel and F.J. Nogales. Portfolio selection with robust estimation. 2007.

[12] L. Garlappi, R. Uppal, and T. Wang. Portfolio selection with parameter and model uncertainty: A multi-prior approach. *Review of Financial Studies*, 20:41–81, 2007.

[13] L. El Ghaoui, M. Oks, and F. Oustry. Worst-case value-at-risk and robust portfolio optimization: A conic programming approach. *Operations Research*, 51(4):543–556, 2003.

[14] D. Goldfarb and G. Iyengar. Robust portfolio selection problems. *Mathematics of Operations Research*, 28(1):1–38, 2003.

[15] B. V. Halldórsson and R. H. Tütüncü. An interior-point method for a class of saddle-point problems. *Journal of Optimization Theory and Applications*, 116(3):559–590, 2003.

[16] R. Jagannathan and T. Ma. Risk reduction in large portfolios: Why imposing the wrong constraints helps. *Journal of Finance*, 58(4):1651–1684, 2003.

[17] R. Johnson and D. W. Wichern. *Applied Multivariate Statistical Analysis*. Prentice Hall, Englewood Cliffs, NJ, 1992.

[18] P. Jorion. Bayes-stein estimation for portfolio analysis. *Journal of Financial and Quantitative Analysis*, 21(3):279–92, 1986.

[19] P. Krokhmal, J. Palmquist, and S. Uryasev. Portfolio optimization with conditional value-at-risk objective and constraints. Research Report 99-14, Department of Industrial and Systems Engineering, University of Florida, 2001.

[20] M.S. Lobo and S. Boyd. The worst-case risk of a portfolio. Information Systems Laboratory, Stanford University, 1999.

[21] H. Markowitz. Portfolio selection. *Journal of Finance*, 7:77–91, 1952.

[22] H. Markowitz. The optimization of a quadratic function subject to linear constraints. *Naval Res. Logistics Q.*, 3:111–33, 1956.

[23] G. Marsaglia. Choosing a point from the surface of a sphere. *Ann. Math. Stat.*, 43: 645–646, 1972.

[24] R.C. Merton. On estimating the expected return on the market: An exploratory investigation. *Journal of Financial Economics*, 8:323–361, 1980.

[25] R. Michaud. The markowitz optimization enigma: is 'optimized' optimal? *Financial Analysts Journal*, 43(1):31–42, 1989.

[26] R. O. Michaud. *Efficient Asset Management*. Harvard Business School Press, Boston, 1998.

[27] M. E. Muller. A note on a method for generating points uniformly on n-dimensional spheres. *Comm. Assoc. Comput. Mach.*, 2:19–20, 1995.

[28] C. Perret-Gentil and M.-P. Victoria-Feser. Robust mean-variance portfolio selection. FAME Research paper no 140, 2004.

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

[30] R. H. Tütüncü and M. Koenig. Robust asset allocation. *Annals of Operations Research*, 132(1):157–187, 2004.

[31] B. Vaz-de Melo and R. P. Camara. Robust modeling of multivariate financial data. Coppead Working Paper Series No. 355, 2003.

[32] Ľ. Pástor and R. F. Stambaugh. Comparing asset pricing models: An investment perspective. *Journal of Financial Economics*, 56:335–381, 2000.

[33] L. Zhu, T. F. Coleman, and Yuying Li. Min-max robust and cvar robust mean-variance portfolios. 2007.

[34] S.S. Zhu and M. Fukushima. Worst-case conditional value-at-risk with application to robust portfolio management. Technical Report 2005-6, Department of Applied Mathematics and Physics, Kyoto University, 2006.