Robust Portfolio Optimization

by

I-Chen Lu

A thesis submitted to
The University of Birmingham
for the degree of
Master of Philosophy (Sc, Qual)

School of Mathematics
The University of Birmingham
July 2009

UNIVERSITY OF BIRMINGHAM

University of Birmingham Research Archive

e-theses repository

ABSTRACT

The Markowitz mean-variance portfolio optimization is a well known and also widely used investment theory in allocating the assets. However, this theory is also familiar with the extremely sensitive outcome by the small changes in the data. Ben-Tal and Nemirovski [3] therefore introduced the robust counterpart approach of the optimization problem to provide more conservative results. And on the ground of their work, Schöttle [26] furthermore proposed the local robust counterpart approach with the smaller uncertainty set.

This paper presents an overview of the local robust counterpart approach of the optimization problem with uncertainty. The classical mean-variance portfolio optimization problem is presented in the first place, and followed by the description of the general convex conic optimization problem with data uncertainty. Afterwards, the concept of the local robust counterpart approach of the optimization problem will be discussed and then applied into the foreign currency market.

Acknowledgements

I would like to acknowledge and extend my heartfelt gratitude to all those who made this dissertation possible.

First of all, I would like to express my sincere gratitude to my advisor Dr Jan Rückmann who offered me the opportunity to write a dissertation about the robust portfolio optimization. His wide knowledge on mathematics and logical way of thinking were of invaluable assistance for me. From him, I have learned not only the research ability, but also the conscientious and careful attitude towards research.

I am also deeply grateful to all the professors and staffs of the School of Mathematics at The University of Birmingham for their instruction, teaching and help. In addition, I wish to extend my heartfelt thanks to my course mates and friends who gave constant support and comfort.

Last but not least, my deepest appreciation goes to my parents and my brother. Their unconditional love, endless support and encouragement enabled me to finish this paper. Thank you.

CONTENTS

Notation and Symbols

$\emptyset$	The empty set
$\stackrel{\mathrm{d}}{=}$	The equality of the distributions
$d(x, y)$	The distance from a point x to another point y
$e \in \mathbb{R}^n$	The column vector defined as $e = (1, \cdots, 1)^T$
H-lsc	The abbreviation of Hausdorff lower semicontinuous
H-usc	The abbreviation of Hausdorff upper semicontinuous
$\inf X$	The infimum of the set X
$\mathrm{int} X$	The interior of the set X
$\wp(X)$	The power set of the set X
$\mathbb{R}^n$	The real $n-$space
$\mathbb{R}_+$	The nonnegative orthant of $\mathbb{R}$
$V_\varepsilon(x)$	The ε-neighborhood of a point x
$V_\varepsilon(X)$	The ε-neighborhood of a set X
$f : \mathbb{R}^n \to \mathbb{R}$	The objective function from $\mathbb{R}^n$ to $\mathbb{R}$
$f^* : U \to \mathbb{R}$	The optimal value function
$g : \mathbb{R}^n \to \mathbb{R}$	The constraint function from $\mathbb{R}^n$ to $\mathbb{R}$
E	The expected return of the portfolio
F	The feasible set of the original optimization problem
F^*	The optimal set of the original optimization problem
F^S	The set of Slater points of the original optimization problem
F_U	The set of all feasible solutions for every uncertain parameter u of the original optimization problem
F_{LRC}	The feasible set of the local robust counterpart optimization problem
F_{LRC}^*	The optimal set of the local robust counterpart optimization problem
F_{LRC}^S	The set of Slater points of the local robust counterpart optimization problem
μ	The expected returns of the assets
ρ	The correlation
$\Sigma \in \mathbb{R}^{n \times n}$	The covariance matrix of the asset returns
σ	The standard deviation
u	The uncertain parameter
U	The uncertainty set
V	The variance of the return of portfolio
ϱ	The risk aversion coefficient
x	The weights in a portfolio
x^*	The optimal solution
X	The feasible set

CHAPTER 1

INTRODUCTION

The mean variance optimization problem developed by Markowitz is one of the leading investment theories in finance. This theory asserts that the return and risk are the most important two parameters for determining the investment allocation, and on the basis of the trade-off between the return and risk, the optimal portfolio can be then obtained. Nevertheless, since the mean variance optimization problem is very sensitive to the perturbations of the input data, therefore, the optimal portfolio generated by this optimization problem is not vary reliable if the incorrect parameter values are adopted.

There are various discussions and investigations on how to eliminate or at least decrease the possibility of using the incorrect input data in the optimization problem. Many papers propose to use robust estimators as input parameters, some others suggest to reformulate the optimization problem into the stochastic programming, and a more different approach to robustification of the optimization problem is to use the idea of resampling which proposed by Michaud [22] in 1998.

Moreover, Ben-Tal and Nemirovski [3] introduced the other idea of robustification called the robust counterpart approach by optimizing over an uncertainty set that based on the incorrect or uncertain data. This specific approach has gained lots interest since it was established, and many papers were based on the studies of determining the uncertainty set of the robust counterpart approach.

The main purpose of this dissertation is to study and investigate the robust counterpart

approach and then apply this process in the foreign currency market.

1.1   Thesis organization

In Chapter 2, a brief introduction of the basic concept of the Markowitz mean variance optimization problem will be introduced. In addition, since there are several different but equivalent formulations of the mean variance optimization problem, therefore, the most widely used formulations will be presented. The main part of chapter 2 is to show the fundamental weakness of the mean variance portfolio optimization problem by giving a simple example in the foreign currency market.

Chapter 3 focuses on the general descriptions of the convex conic optimization problem with the uncertainty, and the idea of the uncertainty set will be first introduced at this point. Moreover, the stability statements of the convex conic optimization problem will be provided as well. The notion of the (local) robust counterpart and the existence of the Slater point for the robustified optimization problem will be given in chapter 4.

Finally, the robust portfolio optimization by applying the (local) robust counterpart approach to the general convex conic optimization problem is presented in Chapter 5 with the example of the foreign currency market. And in order to facilitate the creation of the ellipsoid uncertainty set for the robust portfolio optimization problem, the concepts of statistical distributions for modeling the underlying market and parameter estimators for the inputs will be explained beforehand.

Chapter 2

The mean variance optimization

In the early 1950s, Harry Markowitz introduced the theory of portfolio selection, which is the fundamental idea of the mean variance optimization, the modern portfolio theory, and the mean-variance analysis. The theory of portfolio selection suggests investors that a good portfolio is not just a collection of many good stocks and bonds, but it should also consider the risk and return of the investment according to the investors' objective [20].

The mean variance optimization is also proposed by Harry Markowitz, and this is one of the most common quantitative methodologies used for allocating assets. The main idea of the mean variance optimization is to use the expected return of the assets and the variance of the expected return of assets as the expected return and risk of the investment respectively. The outcome of this methodology is the efficient frontier which offers investors the set of optimal portfolios. By using the mean variance optimization model, the investors can benefit from choosing either the portfolio with the specific expected return and the minimum risk of the investment or the portfolio with the specific level of risk for the investment but the maximum return on the investment.

However, the mean variance optimization problem is very sensitive with the input data: the expected returns and the variances. A slightly changes of the expected returns will lead to a different arrangement in assets allocation. Therefore, for the past few decades, many scholars have proposed different point of views in order to provide a more accurate and stable model to allocate the assets.

This chapter will review the relevant underpinning theories and arguments based on Cornuejols and Tütüncü [10], Elton, Gruber, Brown and Goetzmann [11], and Markowitz [21]. Furthermore, it will also address different scholars' thoughts within the field of management mathematics in the current literature.

2.1 The standard model for mean variance optimization

Before starting the description of the mean variance optimization problem, there are some assumptions that have to be addressed.

First of all, the markets are perfectly efficient without taxes and transaction costs, and the investors are assumed to be all risk averse, which means that the investors will only agree to take higher risk if the higher expected return is offered. In addition, the investors also prefer higher return than lower and less risk than more. Finally, the model supposes that the investment decision would only based on the information of the expected returns and risks.

Consider the model that contains n assets $S_1, S_2, \ldots, S_n$ with weights $x_1, x_2, \ldots, x_n$ invested in those n assets respectively, where

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

$$x_i \geq 0 \qquad \text{with} \qquad i = 1, \ldots, n \tag{2.2}$$

And suppose each individual assets $S_1, S_2, \ldots, S_n$ have expected return $\mu_1, \mu_2, \ldots, \mu_n$, therefore, the expected return of the portfolio E can be stated as:

$$E = \sum_{i=1}^{n} x_i \mu_i \tag{2.3}$$

The risk of the investment can be denoted as the variance of the return $\sigma_1^2, \sigma_2^2, \ldots, \sigma_n^2$,

and σ_{ij} as the covariance between asset i and asset j, where

$$\sigma_{ii} = \sigma_i^2 \tag{2.4}$$

$$\sigma_{ij} = \rho_{ij}\sigma_i\sigma_j \tag{2.5}$$

and ρ_{ij} is the correlation between asset i and asset j. Hence, the variance of the return of portfolio V can be stated as below:

$$V = \sum_{i,j=1}^{n} \rho_{ij}\sigma_i\sigma_j x_i x_j = \sum_{i,j=1}^{n} \sigma_{ij} x_i x_j \tag{2.6}$$

By considering the formulas in the form of matrix, the weights of the portfolio w, the expected returns u, and the covariance between assets $\Sigma = (\sigma_{ij})$ (a $n{\times}n$ symmetric matrix with $\sigma_{ij} = \sigma_{ji}$ and $\sigma_{ii} = \sigma_i^2$) can be written as followed:

$$x = (x_1, \ldots, x_n)^T \tag{2.7}$$

$$\mu = (\mu_1, \ldots, \mu_n)^T \tag{2.8}$$

$$\Sigma = (\sigma_{ij}) = \begin{pmatrix} \sigma_{11} & \cdots & \sigma_{1n} \\ \vdots & \ddots & \vdots \\ \sigma_{n1} & \cdots & \sigma_{nn} \end{pmatrix} \tag{2.9}$$

As the result, the expected return of the portfolio E and the variance of return on the portfolio V can be rearranged as

$$E = \mu^T x \tag{2.10}$$

$$V = x^T \Sigma x \tag{2.11}$$

Definition 2.12 *A portfolio in the standard model is feasible if it satisfies the constraint equations (2.1) and (2.2). In addition, the feasible set is the collection of all feasible portfolios.*

Definition 2.13 *A feasible portfolio is efficient if it has the maximal expected return among all other feasible portfolios that have the same level of risk (the same variance), and has the minimum risk among all other feasible portfolios that have the same expected return.*

Definition 2.14 *The efficient frontier is the set of all efficient feasible portfolios.*

Definition 2.15 *A feasible portfolio is the global minimum variance portfolio (GMV) if it has the lowest level of risk (the lowest variance).*

Figure 2.16 The rough graphical example of the feasible set[1]
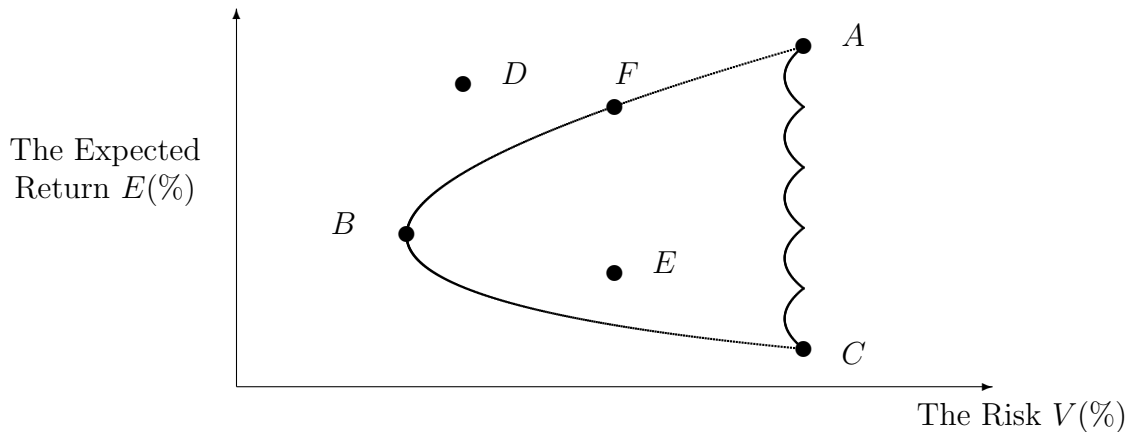
In Figure 2.16, the feasible set is the area that bounded by the curve ABC, and all other portfolios which is outside of the area ABC are the unobtainable portfolios. For instance, the portfolio D is the portfolio which is not feasible.

The portfolios that belong to the area ABC are all feasible but not all efficient. Except the portfolios on the arc AB, the rest of the feasible portfolios can be either replaced by the portfolio with higher expected return but the same level of risk or the portfolio with the same amount of expected return but less variance. The set of all efficient portfolio is called the efficient frontier which is the arc AB in Figure 2.16, and the replaceable portfolios are referred as the inefficient portfolios.

---

[1]Figure 2.16 is just a simple sketch that based on Exhibit 2.1 in Fabozzi, Kolm, Pachamanova, and Focardi [12].

As shown in the Figure 2.16, both portfolios E and F are feasible, but for taking the same level of the risk, the portfolio F can offer more amount of return in the future. Therefore, the investors will rather choose the portfolio F instead of the inefficient portfolio E.

Finally, the portfolio B is the feasible portfolio with the lowest level of risk. Therefore, the portfolio B is the global minimum variance portfolio (GMV).

2.2 Alternative formulations of the mean variance optimization

There are few alternative formulations of the mean variance optimization problem, and the most practical three substitutions are the risk minimization formulation, the expected return maximization formulation and the risk aversion formulation. Those formulations are different as they have dissimilar investment targets, but on the other hand, they are equivalent due to the same efficient frontier which is created by using the expected return and the variance of the portfolio in the similar manner.

The risk minimization formulation is helpful for the investment which requires a minimum variance portfolio with a targeted amount of expected return μ_0, and the formulation is a quadratic optimization problem[2] that can be defined as:

$$\min_{x} \quad x^T \Sigma x \tag{2.17}$$
$$\mu^T x \quad = \quad \mu_0$$
$$e^T x \quad = \quad 1$$

The expected return maximization formulation is applied when the investment has to be kept at a specific level of risk. Hence the optimization problem is designed to maximize

---

[2]Basically, the quadratic optimization problem is a mathematical optimization that optimizing quadratic function subjects to linear constraints. Please refer to Cornuejols and Tütüncü [10] chapter 7 for detailed explanation.

the expected return under the restriction of keeping risk at certain level σ_0^2.

$$\max_{x} \quad x^T \mu \tag{2.18}$$
$$x^T \Sigma x \quad = \quad \sigma_0^2$$
$$e^T x \quad = \quad 1$$

The risk aversion formulation is based on the consideration of the trade-off between the expected return and the risk of the portfolio by introducing the risk aversion coefficient ϱ. This coefficient is also named as the Arrow-Pratt risk aversion index and usually settled between 2 to 4 for the purpose of allocating assets in the investment management.

$$\max_{x} \quad (x^T \mu - \varrho x^T \Sigma x) \tag{2.19}$$
$$e^T x \quad = \quad 1$$

If the aversion to risk is low, then the coefficient ϱ is small and leads to the results of more risky portfolios with higher expected return. In the same way, if the risk aversion is high, then the coefficient ϱ will be large and the optimization problem will result in the portfolios with less risk and lower expected return.

2.3 The example in the foreign currency market

In the market of the foreign currencies, the exchange rate[3] is referred as the price of the foreign currency in the domestic currency. In other words, the exchange rate is the price that the investors need to pay in the sterling pounds for one unit of the other currency, say USD $ 1.

Assuming there is a portfolio with only two foreign currencies, the Euro S_e and the

---

[3]The exchange rates in this paper are spot rates, unless specified otherwise.

US dollar S_{us}, and the domestic currency is the sterling pound. The time period of the investment is one month started from 1st of June.

By collecting the historical data from the website of The Bank of England[4], the exchange rates of the Euro and the US dollar are plotted in the figure below.

Figure 2.20
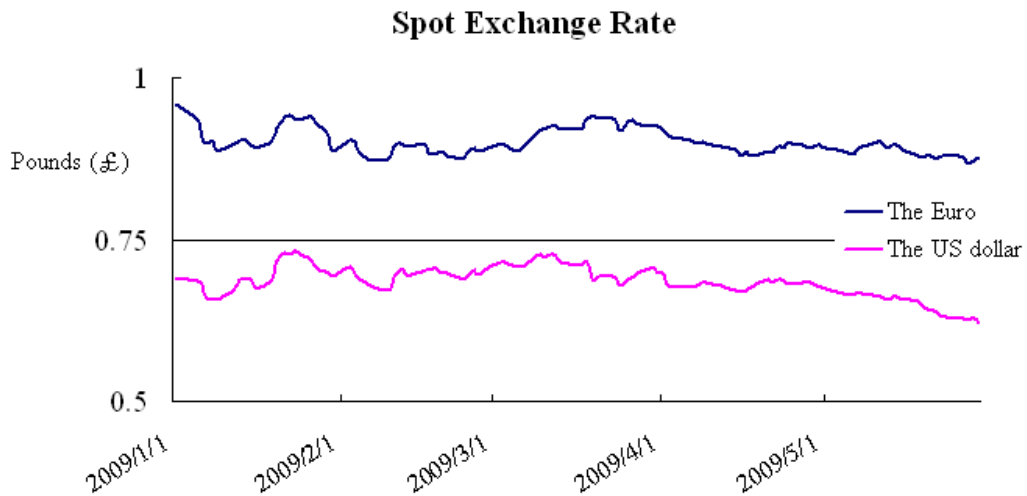
As presented in the figure 2.20, the behavior of the Euro is quite similar to the US dollar, and this fact indicates that these two foreign currencies are highly correlated to each other. But since the portfolio optimization considers the expected return as one of the inputs. Therefore, instead of using the exchange rate as the inputs directly, the monthly rate of return[5] that based on the exchange rates is more appropriate to adopt in this case.

---

[4] The source of the historical data is from the website http://www.bankofengland.co.uk/statistics/index.htm.

[5] The single-period rate of return $= \frac{V_f - V_i}{V_i}$ with V_f and V_i defined as the final value and the initial value of the investment respectively.

Figure 2.21
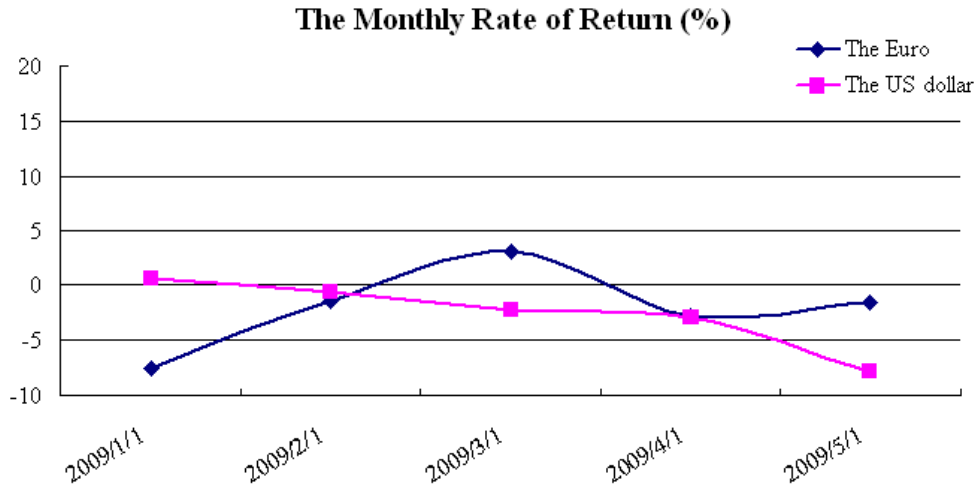
Unlike the resembling performance of the exchange rates in figure 2.20, the monthly rate of return of the Euro behaves differently with the monthly rate of return of the USA dollar. And it can be easily observed in figure 2.21, the swing of the Euro is more aggressive compared to the US dollar as the figure of the Euro moves in a wider range.

Table 2.22 The monthly rate of return from January 2009 - May 2009

Month	EUR (%)	USD (%)
Jan	−7.493	0.595
Feb	−1.471	−0.652
Mar	3.07	−2.269
Apr	−2.815	−2.962
May	−1.581	−7.859

First of all, the expected return of the portfolio is $\mu = (\mu_e, \mu_{us})$, where

$$\mu_e = \frac{-7.493\% - 1.471\% + 3.07\% - 2.815\% - 1.581\%}{5} = -2.058\%$$

$$\mu_{us} = \frac{0.595\% - 0.652\% - 2.269\% - 2.962\% - 7.859\%}{5} = -2.63\%$$

10

The variance $\sigma^2 = (\sigma_e^2, \sigma_{us}^2)$ and the covariance Σ between the Euro and the US dollar are

$$\sigma_e^2 = 11.397\% \qquad \text{and} \qquad \sigma_{us}^2 = 8.379\%$$

$$\Sigma = \begin{pmatrix} \sigma_e^2 & \sigma_{eus} \\ \sigma_{eus} & \sigma_{us}^2 \end{pmatrix} = \begin{pmatrix} 11.397\% & -3.351\% \\ -3.351\% & 8.379\% \end{pmatrix}$$

Therefore, with the weights $x = (x_e, x_{us})^T$, the expected return of the portfolio E and the variance of the portfolio V are

$$
\begin{aligned}
E &= \mu^T x \\
&= -2.058 x_e \% - 2.63 x_{us} \% \\
V &= x^T \Sigma x \\
&= \begin{pmatrix} x_e & x_{us} \end{pmatrix} \begin{pmatrix} 11.397\% & -3.351\% \\ -3.351\% & 8.379\% \end{pmatrix} \begin{pmatrix} x_e \\ x_{us} \end{pmatrix} \\
&= 11.397 x_e^2 \% + 8.379 x_{us}^2 \% - 6.702 x_e x_{us} \%
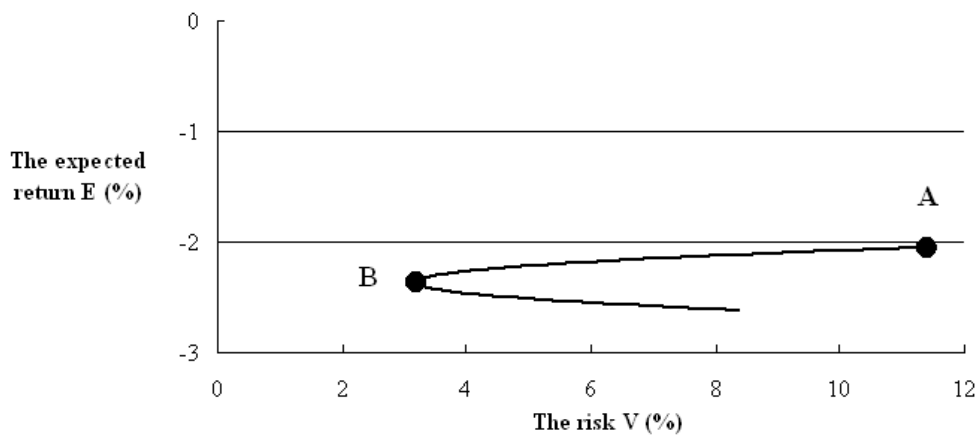\end{aligned}
$$

Figure 2.23

Figure 2.23 shows the feasible set of the investment in the foreign currencies, and

11

the efficient frontier starts from the global minimum variance portfolio B (44% in Euros and 56% in US dollars) to portfolio A. Although none of the feasible portfolios provides possible profit in the future, the maximum expected return portfolio A offers the least loss by consisting 100% of Euros.

As a matter of the fact, in the single period mean variance optimization problem as the given foreign currency example here, the maximum expected return portfolio is always generated by consisting only the asset with the highest expected return. And it is impossible to make any other portfolio with greater expected return than the highest expected return asset. But in the case of considering the risk, usually it is possible to create a less risky portfolio by selecting between different assets, and this is depending on the covariance between each individual assets.

Furthermore, if using the historical information as the inputs, the more data collected the more accurate the findings will be. Therefore, instead of using the previous inputs, the data from January 2000 to December 2008[6] would be adopted for the comparison example.

Figure 2.24
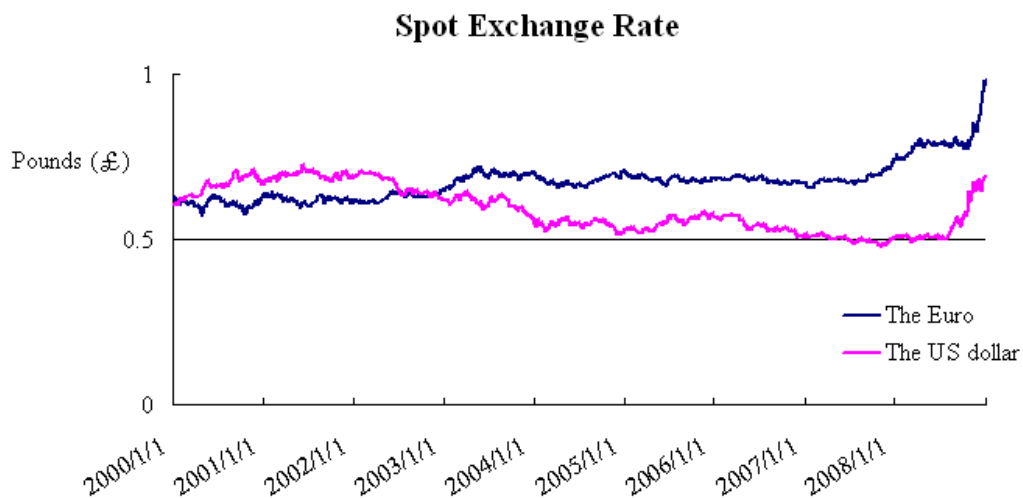
The Euro and the US dollar have approximate same exchange rate to start with, and

---

[6]See Appendix A for the historical data of the monthly rate of return.

the US dollar was stronger than the Euro in the first two years but then started to decline till year 2008. On the contrary, the value of the Euro increases 53% over the past nine years. Figure 2.24 furthermore shows the currencies have opposite performances from year 2000 to year 2003, but since then, the behavior of the currencies become more positive correlated.

Figure 2.25
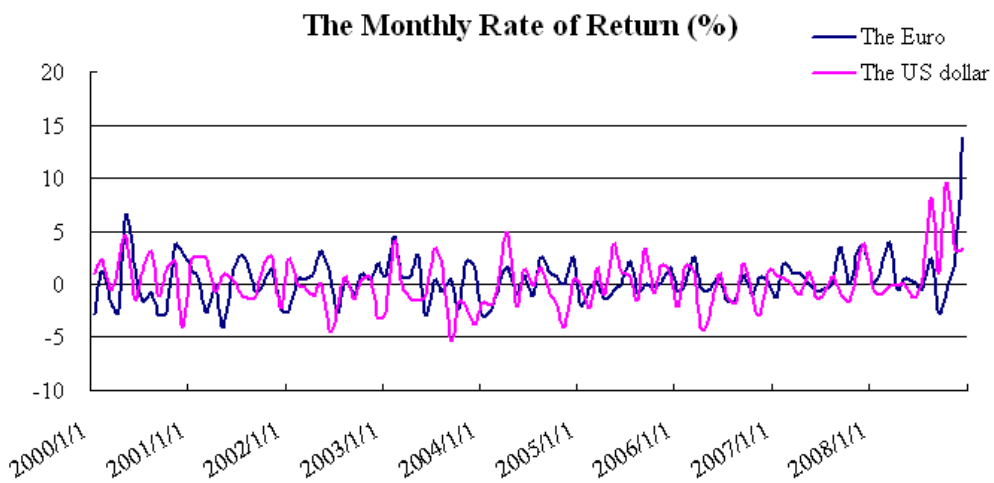
The interesting fact of the figure 2.25 is, although the patterns move drastically during the inspection time period, but generally the movements stay within the range of 5 to -5 except for few instances. Moreover, since the figures are based on the changes of the exchange rate, therefore, the fluctuations are smaller if the performances of the currencies are more stable.

The inputs of the mean variance portfolio optimization in the new sample with data from year 2000 to year 2008 are :

The expected return of each currencies $\mu = (\mu_e, \mu_{us}) = (0.368\%, 0.142\%)$

The variance of each currencies $\sigma^2 = (\sigma_e^2, \sigma_{us}^2) = (5.189\%, 5.714\%)$

The covariance matrix between the currencies

$$\Sigma = \begin{pmatrix} \sigma_e^2 & \sigma_{eus} \\ \sigma_{eus} & \sigma_{us}^2 \end{pmatrix} = \begin{pmatrix} 5.189\% & 1.138\% \\ 1.138\% & 5.714\% \end{pmatrix}$$

With the weights of the portfolio $x = (x_e, x_{us})^T$, the expected return of the portfolio E and the variance of the portfolio V can be estimated as below.

$$
\begin{aligned}
E &= \mu^T x \\
&= 0.368 x_e\% + 0.142 x_{us}\% \\
V &= x^T \Sigma x \\
&= \begin{pmatrix} x_e & x_{us} \end{pmatrix} \begin{pmatrix} 5.189\% & 1.138\% \\ 1.138\% & 5.714\% \end{pmatrix} \begin{pmatrix} x_e \\ x_{us} \end{pmatrix} \\
&= 5.189 x_e^2\% + 5.714 x_{us}^2\% + 2.276 x_e x_{us}\%
\end{aligned}
$$

Figure 2.26
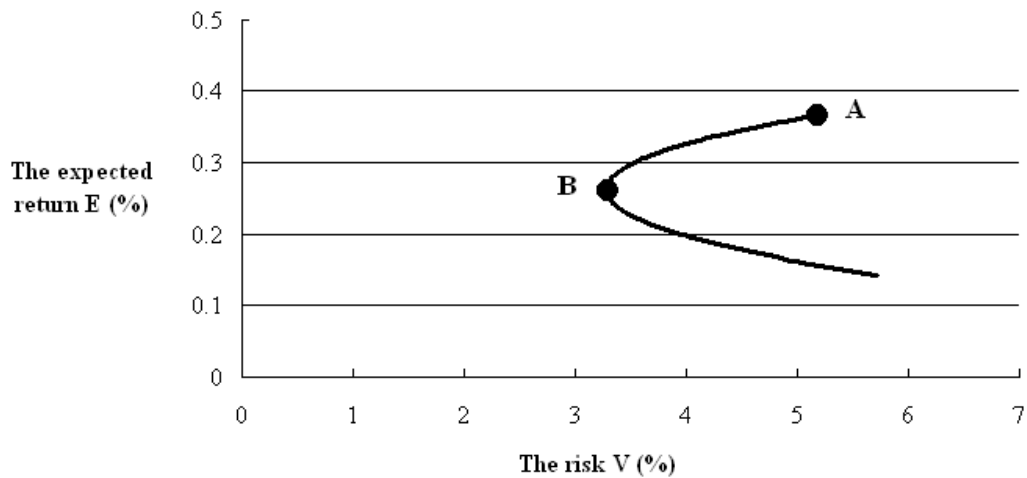
The efficient frontier in Figure 2.26 presents a different result with positive expected return in the future. The curve of the efficient frontier starts from the global minimum variance portfolio B (53% in Euros and 47% in US dollars) to the maximum expected return portfolio A (100% in Euros). Actually, instead of the previous result of no portfolio with possible profit in the future, all of the feasible portfolios are predicted with positive return in the current example.

14

Judging from the above, the inputs of the mean variance portfolio optimization are very important in this methodology. The different sampling data will effect the final investment decision. For instance, if the investor uses the information as the first example, he/she might decide not invest anything at the present and prefer to wait and observe the market movement. But the second example says a different story, and the investor may have the chance to gain some profit.

According to Best and Grauer [5], the changes in the expected return of assets have a great effect on the estimation of efficient portfolios, and this kind of argument has been well studied by several scholars. Frankfurter, Phillips and Seagle [14] used the experimental results to show that the portfolio which selected by the mean variance optimization is unlikely to be as efficient as the equally weighted portfolio, and Britten-Jones [8] showed that the estimations of efficient portfolio weights are extremely sensitive to the estimation errors in the expected return and covariances of the assets.

Instead of concerning the effects of the estimation errors in the expected return, there are some discussions based on the influence of the covariances between assets. Laloux, Cizeau, Bouchaud and Potters [19] have discovered that the mean variance optimization only uses the historical information to determinate the correlation matrix, and the eigenvalues which suppose to control the smallest risk portfolio is dominated by noise. Pafka and Kondor [25] agree with the above finding, and propose to add some constrains on the correlation matrix in order to remove the unnecessary factors and capture the essence of the real correlation structure.

However, some other scholars have suggested another method to solve the issue of the estimation errors, and the main idea of this methodology is to use a set of scenarios instead of a fixed value for a particular parameter. Mulvey, Vanderbei and Zenios [24] used an approach based on the idea of the stochastic programming to generate the solutions from the set of scenarios, and controlled the stability of the solutions by applying the penalty functions. On the other hand, Ben-Tal and Nemirovski [3] considered the uncertain data as the uncertainty set and established the so-called robust counterpart approach to solve

15

the problem with the uncertainty set.

In the later chapters of this paper, the more detailed notion of optimizing with uncertainty will be presented.

CHAPTER 3

DESCRIPTIONS OF THE CONVEX CONIC

OPTIMIZATION WITH THE UNCERTAINTY

SET

The character of a mathematical optimization problem is either minimizing or maximizing an objective function $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ in a feasible set $X \in \mathbb{R}^n$, and usually the feasible set X is restricted by certain constraint functions $g_i : \mathbb{R}^n \longrightarrow \mathbb{R}$, $i = 1, \ldots, m$. Generally, the optimization problem can be formulated in the form:

$$\min_{x} \quad f(x) \tag{3.1}$$
$$subject\ to \quad g_i(x) \leq b_i \quad i = 1, 2, \ldots, m$$

with the optimization variable, a vector $x = (x_1, x_2, \ldots, x_n)$ in $\mathbb{R}^n$, and the limits of the constraint function, constants $b_1, b_2, \ldots, b_m$. The outcome of the optimization problem (3.1) is a vector $x^* \in \mathbb{R}^n$ named the optimal solution, and except it has satisfied the constrain functions g_i, it also has the smallest value among all other possible vectors, that is : $f(x^*) \leq f(y)$ for any vector $y \in \mathbb{R}^n$ that satisfies $g_1(y) \leq b_1, \ldots, g_m(y) \leq b_m$.

There are many different categories of the optimization problems, depending on particular forms of the objective and constraint functions. For instance, the optimization problem (3.1) is a convex optimization problem if the objective function and the con-

straint functions are convex. *i.e.*, $\forall x, y \in \mathbb{R}^n$ and $\alpha \in \mathbb{R}$ with $0 \leqslant \alpha \leqslant 1$

$$
\begin{aligned}
f(\alpha x + (1 - \alpha)y) &\leq \alpha f(x) + (1 - \alpha)f(y) \quad\quad\quad\quad\quad\quad\text{(3.2)} \\
g_i(\alpha x + (1 - \alpha)y) &\leq \alpha g_i(x) + (1 - \alpha)g_i(y) \quad\quad \text{for} \quad i = 1, \ldots, n
\end{aligned}
$$

This chapter is emphasis on presenting the convex conic optimization and the corresponding stability of the optimization problem addressed from [26]. And basically, the convex conic optimization is one type of the mathematical optimization problem that searches the optimal solution of a convex real-valued function defined on a convex cone.

To begin with, some basic definitions and properties of the convex cone will be extracted from the following books : Boyd and Vandenberghe [6] and Brinkhuis and Tikhomirov [7].

3.1   Notations of conic ordering and convexity

Let $K \subset \mathbb{R}^n$ be a nonempty set.

- The set K is convex if : $\forall x, y \in K$, $0 \leq \alpha \leq 1 \Rightarrow \alpha x + (1 - \alpha)y \in K$.

- The set K is a cone if : $\forall x \in K$ and $\forall \lambda \in \mathbb{R}$ with $\lambda \geq 0 \Rightarrow \lambda x \in K$.

- The set K is a convex cone if K is a cone and is also convex[1].

- The cone K is pointed if $K \cap (-K) = \{0_K\}$.[2]

- The cone K is solid if K has nonempty interior[3],and the interior of the cone K is denoted by $\text{int}K$.

- The cone K is an ordering cone (also named the proper cone) if it is closed, convex, pointed, and solid.

---

[1]A cone K is convex if and only if $x + y \in K$ $\forall x, y \in K$

[2]The zero element of the cone K is called the vertex 0_K.

[3]A point $(\bar{k})$ is the interior point of K if $\exists \varepsilon > 0$, such that $V_\varepsilon(\bar{k}) = \{x \in \mathbb{R}^n | \, \|x - \bar{k}\| \leq \varepsilon\} \subset K$.

Definition 3.3 *For each ordering cone $K \subset \mathbb{R}^n$ and $x, y \in \mathbb{R}^n$, the cone ordering $\succeq_K$ on $\mathbb{R}^n$ can be defined by $x \succeq_K y \iff x - y \in K$. In addition, $x \succ_K y \iff x - y \in intK$ defined the strict cone ordering.*

Proposition 3.4 *Let $K \subset \mathbb{R}^n$ be a nonempty proper cone and $\succeq_K$ be the order relation imposed by K in $\mathbb{R}^n$. Then the following conditions are satisfied for any arbitrary $w, x, y, z \in \mathbb{R}^n$.*

- *Reflexivity : $x \succeq_K x$.*

- *Transitivity : $x \succeq_k y$ and $y \succeq_K z \Rightarrow x \succeq_K z$.*

- *Compatibility : $x \succeq_K y$ and $\forall \lambda \in \mathbb{R}$ with $\lambda \geq 0 \Rightarrow \lambda x \succeq_K \lambda y$.*

- *Compatibility : $x \succeq_K y$ and $w \succeq_K z \Rightarrow x + w \succeq_K y + z$.*

- *Antisymmetry : $x \succeq_K y$ and $x \preceq_K y \Rightarrow x = y$.*

Definition 3.5 *Let $K \subset \mathbb{R}^n$ be a proper cone with cone ordering $\preceq_K$, then $f : \mathbb{R}^m \to \mathbb{R}^n$ is K-convex if $\forall x, y \in \mathbb{R}^m$ and $1 \geq \alpha \geq 0$*

$$f(\alpha x + (1 - \alpha)y) \preceq_K \alpha f(x) + (1 - \alpha)f(y). \tag{3.6}$$

The function f is called strictly K-convex if $\forall x, y \in \mathbb{R}^m$, $x \neq y$, and $1 > \alpha > 0$

$$f(\alpha x + (1 - \alpha)y) \prec_K \alpha f(x) + (1 - \alpha)f(y). \tag{3.7}$$

Definition 3.8 *Let $K \subset \mathbb{R}^n$ be a cone, the dual cone of the cone K is the set*

$$K^* = \{y \in \mathbb{R}^n \mid x^T y \geq 0 \quad \forall x \in K\} \tag{3.9}$$

The dual cone K^* of the cone K is a closed convex cone, and has several characteristic properties. For instance, if K has nonempty interior, then K^* is pointed; and if the

closure of K is pointed, then K^* has nonempty interior; furthermore, K^{**} is the closure of the convex hull of K, and the last feature implies that if K is a closed convex cone, then $K^{**} = K$. In addition, a cone is said to be self-dual if $K = K^*$.

Example 3.10 *One of the simple examples to illustrate cone ordering and the dual cone is the set K of nonnegative numbers in $\mathbb{R}$, i.e., $K = \mathbb{R}_+$. Although the set K is only a group of positive numbers on the real line, it still qualifies as a proper cone with the cone ordering $\preceq_K$ performs as the inequality $\leq$. And the dual cone K^* of cone K is*

$$
\begin{aligned}
K^* &= \{y \in \mathbb{R} \mid y^T x \geq 0 \ \forall x \in K\} \\
&= \{y \in \mathbb{R} \mid y^T x \geq 0 \ \forall x \in \mathbb{R}_+\} \\
&= \{y \in \mathbb{R} \mid y \geq 0\} \\
&= \mathbb{R}_+ \\
&= K
\end{aligned}
$$

This example can be extend to the case when the set K is the first orthant in $\mathbb{R}^n$, $K = \mathbb{R}^n_+ = \{x \in \mathbb{R}^n \mid x_i \geq 0 \ i = 1, 2, \ldots, n\}$, then K is a proper cone with the coordinatewise order between vectors, i.e.,

$$
x \geq y \ \forall x, y \in \mathbb{R}^n \Leftrightarrow x_i \geq y_i \ \forall i = 1, \ldots, n.
$$

The dual cone K^* of cone $K = \mathbb{R}^n_+$ is given by

$$
\begin{aligned}
K^* &= \{y \in \mathbb{R}^n \mid y^T x \geq 0 \ \forall x \in K\} \\
&= \{y \in \mathbb{R}^n \mid y^T x \geq 0 \ \forall x \in \mathbb{R}^n_+\} \\
&= \{y \in \mathbb{R}^n \mid y_i \geq 0 \ i = 1, 2, \ldots, n\} \\
&= \mathbb{R}^n_+ \\
&= K
\end{aligned}
$$

Proposition 3.11 *Let $K \subset \mathbb{R}^n$ be a proper cone with cone ordering $\preceq_K$, then the function $f : \mathbb{R}^m \to \mathbb{R}^n$ is K-convex if and only if $z^T f(.)$ is convex $\forall z \succeq_{K^*} 0$.*

Proof. If the function f is K-convex, then $\forall x, y \in \mathbb{R}^m$ and $\alpha \in [0, 1]$

$$f(\alpha x + (1 - \alpha)y) \preceq_K \alpha f(x) + (1 - \alpha)f(y) \tag{3.12}$$

By Definition 3.3, (3.12) is equivalent to

$$\alpha f(x) + (1 - \alpha)f(y) - f(\alpha x + (1 - \alpha)y) \in K \tag{3.13}$$

By Definition 3.8,

$$\forall z \in K^* \qquad (\alpha f(x) + (1 - \alpha)f(y) - f(\alpha x + (1 - \alpha)y))^T z \geq 0$$
$$\Rightarrow \qquad \langle \alpha f(x) + (1 - \alpha)f(y) - f(\alpha x + (1 - \alpha)y), z \rangle \geq 0$$
$$\Rightarrow \quad \langle z, \alpha f(x) \rangle + \langle z, (1 - \alpha)f(y) \rangle - \langle z, f(\alpha x + (1 - \alpha)y) \rangle \geq 0$$
$$\Rightarrow \quad \alpha \langle z, f(x) \rangle + (1 - \alpha)\langle z, f(y) \rangle \geq \langle z, f(\alpha x + (1 - \alpha)y) \rangle$$
$$\Rightarrow \quad \alpha z^T f(x) + (1 - \alpha)z^T f(y) \geq z^T f(\alpha x + (1 - \alpha)y)$$

Therefore, $z^T f$ is convex $\forall z \in K^*$.

The converse proof also follows directly from the same definitions; hence it would be omitted. $\qquad \square$

Moreover, if the cone $K \subset \mathbb{R}^n$ has nonempty interior, then

$$x \in int K \Leftrightarrow y^T x > 0 \ \forall \, y \in K^* \backslash \{0\}. \tag{3.14}$$

3.2 The uncertainty set U

One of the serious concerns in most of the optimization problems is the accuracy of the input data. Sometimes this matter can be as simple as the inputs are unknown at the time the optimization problem must be solved; sometimes the inputs are incorrectly computed or just uncertain. As mentioned in Chapter 2, the result of the optimization problem is very sensitive to those input data. And with different inputs, the optimization problem will lead us to the solution that is far from the optimal of the original problem.

The uncertainty of the inputs can be simply modeled by introducing an uncertainty set U that contains many possible values of required uncertain parameters. And there is a corresponding constraint for each uncertain parameter $u \in U$ to ensure the original optimization constraint is fulfilled for the worst-case of the uncertainty set U.

In general, the shape of the uncertainty set U depends on the sources of uncertainty and also the sensitive affection of the uncertainty. The most common shapes of the uncertainty set are interval, ellipsoid and the intersection of the ellipsoids. Schöttle [26] investigated the effects of interval and ellipsoidal uncertainty set on the continuity properties of the optimal solution set. The result illustrated that under certain conditions, the ellipsoidal uncertainty set smoothen the optimal value function, and in most practical cases the ellipsoidal uncertainty set lead to a unique optimal solution that is continuous with respect to the uncertain parameter.

The size of the uncertainty set U depends on the desired robust level of the optimization problem. Some scholars mentioned that the price for the optimization problem with uncertainty set to be stable can be measured in the term of the increase of the optimal value from the original one, and the amount that rose in the optimal value is linear in the size of the uncertainty set. Ben-Tal and Nemirovski [3] used the linear problems to prove the above statement, and Schöttle [26] obtained the same result by examine the convex conic problems.

In addition, Calafiore and Campi [9] proposed to use the scenario approach that based

on the constraint sampling to deal the uncertainty of the optimization problem. The uncertainty set of this methodology is defined as a collection of scenarios for the uncertain parameters, and there is one constraint for each scenario. Therefore, the outcome of the modified optimization problem needs to be qualified for all the constraints that generated by the corresponding scenario. And the number of the required scenarios for the portfolio with n assets can be calculated by fixing the desired level parameter[4] $\varepsilon \in [0, 1]$ and confidence parameter $\beta \in [0, 1]$:

$$\text{The number of the scenarios} \geq \frac{n}{\varepsilon \beta} - 1$$

3.3 The general convex conic optimization problem with data uncertainty (GCP_u)

This section presents the setting for the convex conic optimization problem that depends on the particular uncertain parameter. And for easy reference, some notations and assumptions that will be used throughout the article will be introduced beforehand:

- u, the vector that represent the uncertain parameter.

- x, the vector that represent the decision parameter.

- $U \subset \mathbb{R}^d$, the nonempty convex and compact set of the uncertain parameters.

- $X \subset \mathbb{R}^n$, the nonempty convex and compact set of decisions that do not depend on the uncertain parameter u.

- $K \subset \mathbb{R}^m$, the ordering cone, i.e., K is closed, solid, convex and pointed.

- $f : \mathbb{R}^n \times \mathbb{R}^d \to \mathbb{R}$, the objective function that is continuous in x and u, and f is convex in u for fixed $x \in X$, and convex in x for fixed $u \in U$.

---

[4]The accepted possibility that the original constraint may violated throughout the process of the optimization

- $g : \mathbb{R}^n \times \mathbb{R}^d \to \mathbb{R}^m$, the constraint function that is continuous in x and u, and g is $K-$convex in u for fixed $x \in X$, and $K-$convex in x for fixed $u \in U$.

Definition 3.16 *The general convex conic optimization that based on the uncertain parameter u, (GCP_u), can be formulated as*

$$\min_{x \in X} \quad f(x, u) \tag{3.17}$$
$$subject\ to \quad g(x, u) \preceq_K 0$$

Definition 3.18 *For a given $u \in U$ in the general convex conic optimization problem, the feasible set, the optimal set, and the set of Slater points can be stated as :*

- *The feasible set $F(u) : \{x \in X \mid g(x, u) \preceq_K 0\}$.*

- *The optimal set $F^*(u) : \{x \in F(u) \mid f(x, u) \leq f^*(u)\}$.*

- *The set of Slater points $F^S(u) : \{x \in X \mid g(x, u) \prec_K 0\}$.*

where $f^* : U \to \mathbb{R}$ is the optimal value function defined as

$$f^*(u) := \min\{f(x, u) \mid x \in F(u)\}, \tag{3.19}$$

and the set of all feasible solutions for every $u \in U$ is $F_U = \bigcap_{u \in U} F(u)$.

Additionally, for each $u \in U$ in the general convex conic optimization, the functions that assign the feasible set $F(u)$ and the optimal set $F^*(u)$ are

- The feasible set mapping $F : U \to \wp(X)$.

- The optimal set mapping $F^* : U \to \wp(X)$.

where $\wp(X)$ is the power set of the set X.

Finally, throughout this paper, the set of the solutions that are feasible for all $u \in U$ is assumed to be nonempty unless in the particular situation.

Assumption 3.20 *There is at least one feasible solution $x \in X$ for all parameters within the uncertainty set U, $F_U \neq \emptyset \ \forall u \in U$.*

3.4 The stability of the (GCP_u) optimization problem

In the previous sections, we were only concerned about the setting of the (GCP_u) optimization problem without paying any attention to the corresponding stability of the optimization problem. Generally speaking, a solution of the optimization problem is stable if the solution is not effected too much by the small perturbations of the input data. And the stability property of the optimization problem can be expressed as the continuous mapping[5] between the input data and the solutions of the optimization problem.

Definition 3.21 *A mapping $h : X \rightarrow Y$ is continuous for some $x \in X$ if for any neighborhood V of $h(x)$ there is a neighborhood U of x such that $h(U) \subset V$.*

One of the important requirements for examining the stability (continuity) of the optimization problem is the existence of at least one solution for the optimization problem, and this requirement can be assured by the Assumption 3.20, $F_U \neq \emptyset$, together with the objective function f being continuous. Consequently, the optimal solution set $F^*(u)$ of the (GCP_u) optimization problem is nonempty for every $u \in U$.

Schöttle [26] derived the stability of the (GCP_u) optimization problem by applying the concepts of the Hausdorff continuity and the Berge continuity. In addition, Schöttle proposed that with respect to the parameter $\widehat{u} \in U$, the appearance of the Slater point guarantees the Hausdorff upper semi-continuous property of the feasible set mapping F even under the situation of more than one solution. And the Slater point can be obtained if the constraint function $g(\cdot, \widehat{u})$ of the (GCP_u) optimization problem is strictly $K-$convex. That is, suppose $x, y \in F(\widehat{u})$ are any two solutions of $\widehat{u} \in U$, and since $F(\widehat{u})$ is a convex

---

[5]The ideas of the continuous mapping (function) are:

- For a function $f(.)$, the value of $f(x)$, where x is a certain point belongs to the domain of $f(.)$, can be predicted by using the value of $f(.)$ at some points near x.

- The graph of the function $f(.)$ is a connected curve without any jumps, holes and gaps.

set, then for all $\alpha \in [0, 1]$ a point $z := \alpha x + (1 - \alpha)y$ also belongs to the set $F(\widehat{u})$, such that

$$g(z, \widehat{u}) = g(\alpha x + (1 - \alpha)y, \widehat{u}). \tag{3.22}$$

If the constraint function $g(\cdot, \widehat{u})$ is strictly $K-$convex as required, then the point $z := \alpha x + (1 - \alpha)y$ satisfies the following equation

$$g(\alpha x + (1 - \alpha)y, \widehat{u}) \prec_K \alpha g(x, \widehat{u}) + (1 - \alpha)g(y, \widehat{u}), \tag{3.23}$$

and by the definition of the feasible set, $g(x, \widehat{u}) \preceq_K 0$ and $g(y, \widehat{u}) \preceq_K 0$. Hence, equation (3.23) becomes

$$g(\alpha x + (1 - \alpha)y, \widehat{u}) \prec_K 0 \tag{3.24}$$

which is equivalent to

$$g(z, \widehat{u}) \prec_K 0. \tag{3.25}$$

As the result, the point $z \in F(\widehat{u})$ is a Slater point of the (GCP_u) optimization problem.

Before continuing the investigation about the stability of the (GCP_u) optimization problem, there are some necessary definitions that extracted from the book by Bank [2] need to be addressed.

Definition 3.26 *The $\varepsilon-$neighborhood of a set $S \subset \mathbb{R}^n$ with $\varepsilon > 0$ is defined by*

$$V_\varepsilon(S) := \{x \in \mathbb{R}^n \mid d(x, S) = \inf_{y \in S} d(x, y) < \varepsilon\}. \tag{3.27}$$

Definition 3.28 *A point$-$to$-$set mapping $\Gamma : U \to \wp(\mathbb{R}^n)$ is closed at a point $\widehat{u}$ if the*

sequences

$$u_k \to \widehat{u}, \quad \{u_k\} \subset U$$

$$\text{with } x_k \in \Gamma(u_k), \ \ k = 1, 2, \ldots, \ \ \text{imply } \widehat{x} \in \Gamma(\widehat{u}),$$

$$x_k \to \widehat{x}, \quad \{x_k\} \subset \mathbb{R}^n$$

and the mapping Γ is closed if it is closed at every point of U.

Definition 3.29 *The mapping Γ is Hausdorff upper semi-continuous at the point $\widehat{u}$ if for each $\varepsilon > 0$ there is $\delta > 0$, such that*

$$\Gamma(u) \subset V_\varepsilon(\Gamma(\widehat{u})) \ \ \forall u \in V_\delta(\widehat{u}), \tag{3.30}$$

and the mapping Γ is Hausdorff lower semi-continuous at the point $\widehat{u}$ if for each $\varepsilon > 0$ there is $\delta > 0$, then

$$\Gamma(\widehat{u}) \subset V_\varepsilon(\Gamma(u)) \ \ \forall u \in V_\delta(\widehat{u}). \tag{3.31}$$

Definition 3.32 *The mapping $\Gamma : U \to \wp(\mathbb{R}^n)$ is Hausdorff continuous at the point $\widehat{u}$ if Γ is Hausdorff upper semi-continuous and also Hausdorff lower semi-continuous at the point $\widehat{u}$.*

Remark 3.33 *The mapping $\Gamma : U \to \wp(\mathbb{R}^n)$ is Hausdorff upper semi-continuous at $\widehat{u} \in U$ if Γ is closed and the set X is compact.*

The following graph illustrates a schematic procedure of determining the continuity of the feasible set mapping F with H-usc and H-lsc stand for the abbreviation of the Hausdorff upper semi-continuous and the Hausdorff lower semi-continuous respectively.

Figure 3.34

$$\boxed{\begin{array}{c} \min_x f(x, u) \\ \text{subject to } \ g(x, u) \preceq_K 0 \end{array}}$$

g is continuous and the cone K is closed

$$\boxed{\text{The mapping } F \text{ is closed}}$$

The decision set X is compact

$$\boxed{\text{The mapping } F \text{ is H-usc}}$$

$\exists$ More than one solution and $g \prec_K 0$

$\exists$ unique solution

$$\boxed{\exists \text{ Slater points}}$$

$$\boxed{\text{The mapping } F \text{ is H-lsc}}$$

$$\boxed{\text{The mapping } F \text{ is H-lsc}}$$

$$\boxed{\text{The mapping } F \text{ is Hausdorff continuous at } \widehat{u}}$$
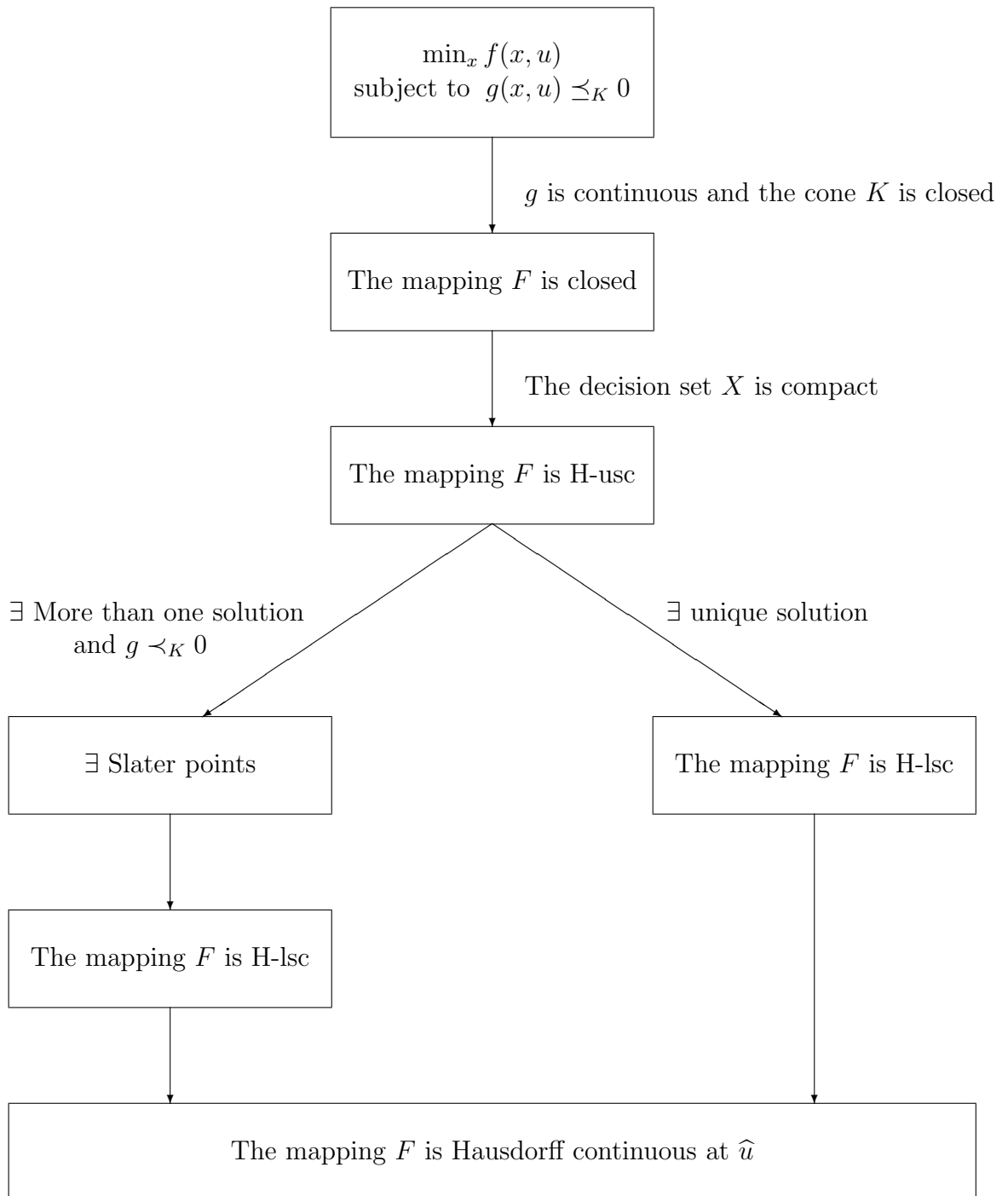
First of all, the feasible set mapping F is closed under the assumptions of the constraint function g is continuous and the cone K is closed. Consequently, the closedness of the mapping F together with the set X being compact verify that the mapping F is Hausdorff upper semi-continuous for all $u \in U$.

Now, according to the assumption $F_U \neq \emptyset$ for all $u \in U$, there is either a unique solution or more than one solutions for each $u \in U$. Hence, instead of describing the Hausdorff lower semi-continuity of the mapping F as one single case, the following two situations will be considered.

- The solution is unique for each $u \in U$.

- There exists more than one solution for each $u \in U$.

In the case of only one solution for each $u \in U$, the Hausdorff upper semi-continuity of the feasible set mapping $F : U \to \wp(X)$ imply the following :

Let $\widehat{u} \in U$, for $\varepsilon > 0$ there exists $\delta > 0$, such that

$$F(u) \subset V_\varepsilon(F(\widehat{u})) \;\; \forall u \in V_\delta(\widehat{u}). \tag{3.35}$$

And because $F(\widehat{u})$ is the only solution for $\widehat{u}$, therefore $F(u) \in V_\varepsilon(F(\widehat{u}))$, where

$$V_\varepsilon(F(\widehat{u})) = \{y \in \mathbb{R}^n \;\mid\; d(y, F(\widehat{u})) < \varepsilon\}. \tag{3.36}$$

Since $F(u) \in V_\varepsilon(F(\widehat{u}))$, then it implies

$$d(z, F(\widehat{u})) < \varepsilon \;\; \forall z \in F(u), \tag{3.37}$$

and hence

$$F(\widehat{u}) \in V_\varepsilon(z) \;\; \forall z \in F(u). \tag{3.38}$$

Which is equivalent to

$$F(\widehat{u}) \in V_\varepsilon(F(u)). \tag{3.39}$$

Therefore, the feasible set mapping F is Hausdorff lower semi-continuous for all $u \in U$ in this case.

Under the circumstance of more than one solution for $\widehat{u} \in U$, the Slater point exists if the constraint function $g(\cdot, \widehat{u})$ is strictly $K-$convex. And according to [26], the feasible set mapping F is Hausdorff upper semi-continuous at $\widehat{u} \in U$ if there is a Slater point of $F(\widehat{u})$.

Eventually, with both Hausdorff upper semi-continuity and Hausdorff lower semi-continuity at $\widehat{u}$, the feasible set mapping F is Hausdorff continuous at $\widehat{u}$.

Chapter 4

The local robust counterpart approach

In 1998, Ben-Tal and Nemirovski published their paper [3] on deriving the robust counterpart approach for the optimization problem with uncertain parameters. This robust counterpart approach is in fact the worst-case approach of the original optimization problem, as the modified optimization problem is not only solved the problem for every point of the uncertain set, but is concerned specially the case with the worst performance. On the ground of this development, Ben-Tal and Nemirovski have continuously investigated the robust counterpart approach in various situations, and one of their recent journals [4] summarized several selected topics of the robust counterpart, especially on the concept of the extended idea and the tractability of the robust counterparts.

Before the introduction of the robust counterpart approach, there are several methodologies to deal with the uncertainty. The most common techniques are the sensitivity analysis and the stochastic programming. In many optimization problems, the uncertain data are replaced by the nominal values and then justified the stability of the results by the sensitivity analysis which only deals with the trivial deviations of the nominal data. On the other hand, the stochastic programming pays attention to the uncertain data from the beginning of the problem, but this method has the following weaknesses: the solution may not satisfy the required constraint and the price of using the stochastic programming

to solve the problem can be very expensive.

Apart from those already mentioned techniques, another approach to solve the optimization problem with uncertainty is the robust mathematical programming developed by Mulvey, Vanderbei and Zenios in 1995 [24]. This programming uses a set of scenarios instead of the fixed nominal data (point estimates), but has the same disadvantage as the stochastic programming. The robust mathematical programming might leads to the solutions that unfulfilled the constraint.

The main motivation of this chapter is to present the local robust counterpart approach together with the stability of the local robust counterpart problem proposed by Schöttle [26]. And to begin with, the concept of the robust counterpart approach by Ben-Tal and Nemirovski [3] would be discussed first.

4.1   The robust counterpart approach

As already mentioned, the robust counterpart approach is the worse -case approach of the original optimization problem with the uncertainty. This methodology first assumes the uncertain parameters may vary in a particular set, and then find the portfolio weights to optimize the objective function even if the "worse" case turns out in the reality.

If the uncertainty set is finite, i.e., $U = \{u_1, u_2, \ldots, u_n\}$, then the optimization problem can be solved by transferring the function that contains the uncertain parameter into finitely many functions for every single $u \in U$, and similar process can be applied if there are finitely many vertices belong to the uncertainty set U. Since the transferring process is only duplicated the particular function, therefore, the original optimization problem is turned into a larger version but not more difficult due to the structural properties of the original optimization problem are preserved. For instance, consider the following general optimization problem with the uncertainty u:

$$\min_{x} \quad f(x, u) \tag{4.1}$$
$$subject\ to \quad g(x, u) \in K$$

If the uncertainty set is a finite set $U = \{u_1, u_2, \ldots, u_n\}$, then the robust counterpart of the general optimization problem (4.1) can be formulated as :

$$\min_{x,z} \quad z \tag{4.2}$$
$$subject\ to \quad z - f(x, u_i) \geqslant 0 \quad i = 1, 2, \ldots, n$$
$$g(x, u_i) \in K \quad i = 1, 2, \ldots, n$$

On the other hand, if the uncertainty set is not finite, i.e., continuous set in the shape of intervals, ellipsoids or the intersections of ellipsoids, then the optimization problem would be modified into a more complicated process as the function that contains the uncertain parameter has to be satisfied for all values in the uncertainty set, and there would be infinitely many constraints needed to be fulfilled in order to solve the function with the uncertain parameter. Hence this converts the original optimization problem into the semi-infinite optimization problem.

The robust counterpart of the (GCP_u) optimization problem, the general convex conic optimization problem with uncertain data u, is a semi-infinite programming, which takes care the problem of the appearance of the worse value of the uncertainty without changing the original features of the (GCP_u) problem, and is formulated as:

$$\min_{x} \max_{u} \quad f(x, u) \tag{4.3}$$
$$subject\ to \quad g(x, u) \preceq_K 0 \quad \forall u \in U$$

where $x \in X$ and $u \in U$. And according to the assumption $F_U \neq \emptyset$ of the (GCP_u) optimization problem, there is a nonempty feasible set for every single $u \in U$. This implies that the robust counterpart optimization problem also has the nonempty feasible set, and hence (4.3) is equivalent to:

$$\min_{x \in F_U} \max_{u \in U} \quad f(x, u) \tag{4.4}$$

4.2 The local robust counterpart approach

This section describes the local robust counterpart approach which based on the robust counterpart approach that mentioned in the previous section. Compared with the robust counterpart approach, the local robust counterpart approach concerns only the smaller area around a particular parameter $\widehat{u}$ of the uncertainty set instead of the entire uncertainty set U, and hence not all of the values in the uncertainty set U require to fulfill the constraint of the local robust counterpart problem.

Definition 4.5 *The local uncertainty set $U_\delta(\widehat{u}) \subset U$ centered at $\widehat{u}$ with a suitable ratio of size $\delta \geqslant 0$ is defined as:*

$$U_\delta(\widehat{u}) = (\widehat{u} + \delta U) \cap U$$

For example, the original uncertainty set U centered at the point u_0 can be displayed in the same format with $\delta = 1$:

$$U = U_\delta(u_0) = (u_0 + \delta U) \cap U = U.$$

Definition 4.6 *For any $\widehat{u} \in U$ and $\delta \geqslant 0$, the pair $(\widehat{u}, \delta)$ is called acceptable if the corresponding set $\widehat{u} + \delta U$ is a subset of the original uncertainty set U.*

In some occasions, the shape of the local uncertainty set $U_\delta(\widehat{u})$ is not elliptical as expected, i.e., some values of $U_\delta(\widehat{u})$ is not within the uncertainty set U. The following figures illustrate 3 examples of the local uncertainty set $U_\delta(\widehat{u})$.

Figure 4.7 The original uncertainty set U
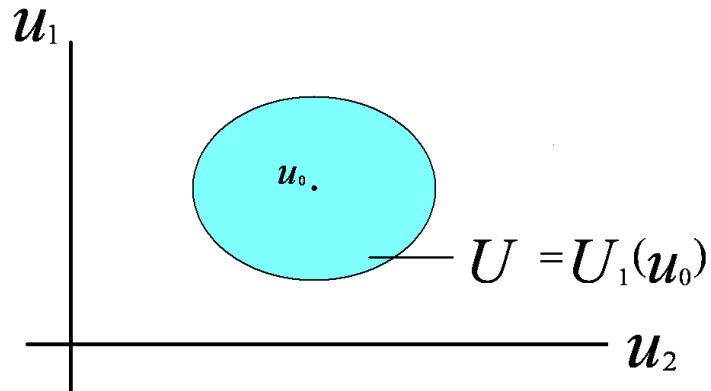

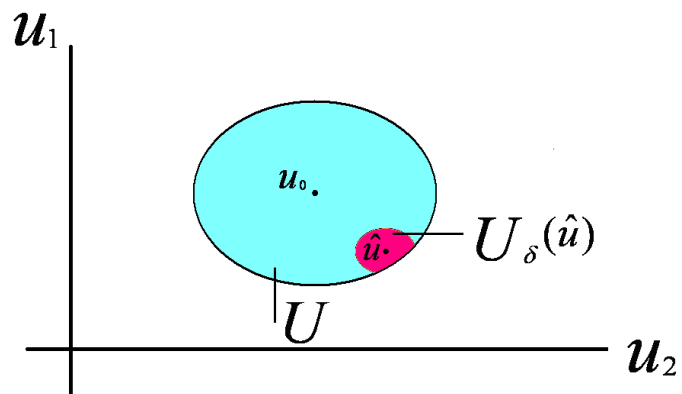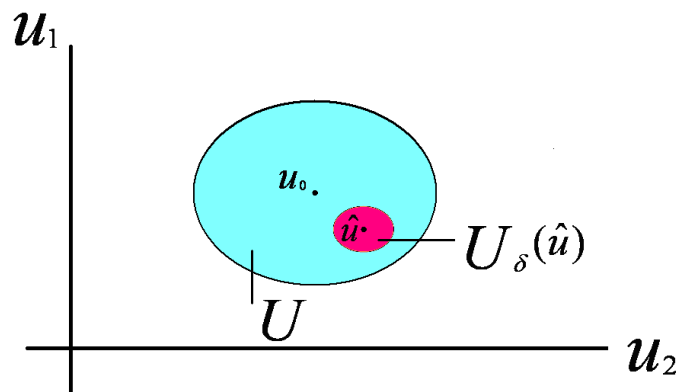**Figure 4.8** The local uncertainty set $U_\delta(\widehat{u})$ with unacceptable pair $(\widehat{u}, \delta)$

Figure 4.9 *The local uncertainty set $U_\delta(\widehat{u})$ with acceptable pair $(\widehat{u}, \delta)$*

Definition 4.10 *The local robust counterpart problem* $(LRC_{\widehat{u},\delta})$ *of the* $(GCP_{\widehat{u}})$ *optimization problem with* $\widehat{u} \in U$ *and* $\delta \geqslant 0$ *be the acceptable pair is formulated as:*

$$\min_{x} \ \max_{u} \quad f(x,u) \tag{4.11}$$

$$subject \ to \quad g(x,u) \preceq_K 0 \quad \forall u \in U_{\delta}(\widehat{u})$$

where $x \in X$ *and* $u \in U_{\delta}(\widehat{u})$.

And under assumption 3.20 ($F_U \neq \emptyset$), the feasible set of the local robust counterpart problem is also nonempty, hence the formulation of the $(LRC_{\widehat{u},\delta})$ problem can be stated in the same way as equation (4.4).

$$\min_{x \in F_{U_{\delta}(\widehat{u})}} \ \max_{u \in U_{\delta}(\widehat{u})} \quad f(x,u) \tag{4.12}$$

The related notations and formulations for the local robust counterpart optimization problem $(LRC_{\widehat{u},\delta})$ would be stated in the Appendix B.

In view of the fact that the existence of a Slater point is an important factor for the stability of the (GCP_u) optimization problem. Therefore, the notion of the Slater point for the local robust counterpart problem will be discussed to close this chapter.

By definition, a point $\widehat{x} \in X$ is a Slater point of the $(LRC_{\widehat{u},\delta})$ problem if

$$g(\widehat{x},u) \prec_K 0 \quad \forall u \in U_{\delta}(\widehat{u}) \tag{4.13}$$

or equivalently,

$$g(\widehat{x},u) \in \text{int}(-K) \quad \forall u \in U_{\delta}(\widehat{u}) \tag{4.14}$$

And if $\widehat{x} \in X$ is a Slater point of the $(LRC_{\widehat{u},\delta})$ problem, then there exists $\varepsilon > 0$ such that

$$V_{\varepsilon}(g(\widehat{x},u)) \subset \text{int}(-K) \quad \forall u \in U_{\delta}(\widehat{u}) \tag{4.15}$$

Schöttle [26] has proposed that the existence of the Slater point of the the (GCP_u) optimization problem implies the existence of the Slater point of the $(LRC_{\widehat{u},\delta})$ problem.

And in addition, there is a global ratio of size $\delta_{glob} > 0$ such that the $(LRC_{\widehat{u},\delta_{glob}})$ problem holds a Slater point for any $\widehat{u} \in \mathrm{int}U$ if the constraint function g is globally Lipschitz continuous[1] for all $u \in U$.

---

[1]A mapping $f : X \to Y$ is Lipschitz continuous on a set $S \subset X$ if for some constant $L \geqslant 0$ and $\forall s_1, s_2 \in S$ the following holds $\| f(s_1) - f(s_2) \| \leqslant L \| s_1 - s_2 \|$.

CHAPTER 5

THE ROBUST PORTFOLIO OPTIMIZATION

The mean variance optimization introduced by Markowitz is also refereed as the portfolio optimization that allows the investors to choose the portfolio with the highest expected return for a given level of risk. This methodology of allocating the assets is famous and widely used in the field of investment. However, as showed in the chapter 2, the outcome of the mean variance optimization is extremely sensitive to the perturbations in the inputs, and hence the subsequent solutions are not very reliable.

There are many discussions on how to decrease or eliminate the possibility of using incorrect inputs for the optimization problem. Some suggest to use the parameter estimators as inputs in order to reduce the sensitivity of the optimal portfolio, and on the other hand, Michaud [22] proposed to use the technique of resampling the input parameters from a confidence region and then average the cumulative portfolios that obtained by each pair of sampling data. The main idea here is, if resampling enough times, then the averaged optimal portfolio should be more stable and less sensitive to the perturbations in the inputs. But this methodology is not efficient when the amount of assets becomes large.

This chapter presents the application of the local robust counterpart approach for the mean variance portfolio optimization problem in the foreign currency market, and this procedure suppose to provide more stable optimal portfolios even under the situation with the uncertain inputs. In the framework of this modified optimization problem, the

uncertain or incorrect parameters are simply modeled by a bounded uncertainty set U, but there is no specific definition or restriction for the uncertainty set U in the modified optimization problem.

In order to define the uncertainty set for the portfolio optimization, one can first assumed the input data follows a multivariate distribution as the model of the financial market, and then choose the parameter estimator for the first two moments of the distribution as the input parameter for the portfolio optimization. The most common and simple ways to model a financial market is to assume the input data of the portfolio optimization follows the normal distribution, but the normal distribution is not always the suitable choice to model the financial market, especially when the return of the underlying assets is violated.

5.1 The statistical distributions for modeling the underlying environment

This section focuses on discussing the statistical distributions for modeling the monthly rate of returns of the foreign currency market that mentioned in Chapter 2.

To begin with, the sample data of the foreign currency market will be considered and investigated to check whether the normal distribution is the suitable choice to model the particular foreign currency market.

5.1.1 Normal distribution

In 1733, Abraham de Moivre introduced the concept of the normal distribution as an approximation for the distribution of the sum of binomial random variables. The normal distribution is a continuous probability distribution with data clustered around the mean μ, and the graph of the associated probability density function is always symmetric and has the bell-shaped curve with a single peak at the mean μ. For instance, the following diagram presents the pattern of the probability density function of the standard normal distribution with mean $\mu = 0$ and variance $\sigma^2 = 1$.

Figure 5.1

The normal distribution
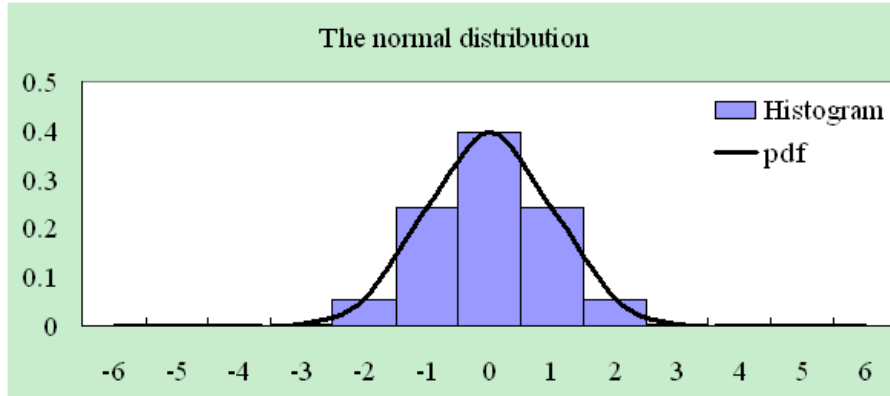
A random variable X with mean μ and variance σ^2 is normally distributed if it has the probability density function as

$$f(x; \mu, \sigma^2) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

where the μ indicates the location of the peak of the density and the variance σ^2 stands for the spread of the bell curve.

The normal distribution describes the sample with the variable that tends to occur around the mean μ, and especially useful for modeling the sample with the large quantity of independent data. Hence, this distribution becomes widely used as a simple model for the complex phenomena in many different fields, such as statistics, natural science, and social science. For a more detailed information of the normal distribution see [1].

According to the historical data of the Euro and the US dollar from January 2000 to December 2008, the expected monthly rate of return (mean) μ and the variance σ^2 of the currencies for this inspected time period are :

Table 5.2

		The Euro	The US dollar
Mean	μ	0.368%	0.142%
Variance	σ^2	5.189%	5.714%

By assuming both the Euro and the US dollar follow the normal distribution with the corresponding expected rate of return μ and variance σ^2, the comparison between the curve of the probability density function and the histogram of the monthly rate of returns of each currency can be then presented separately in the following figures.
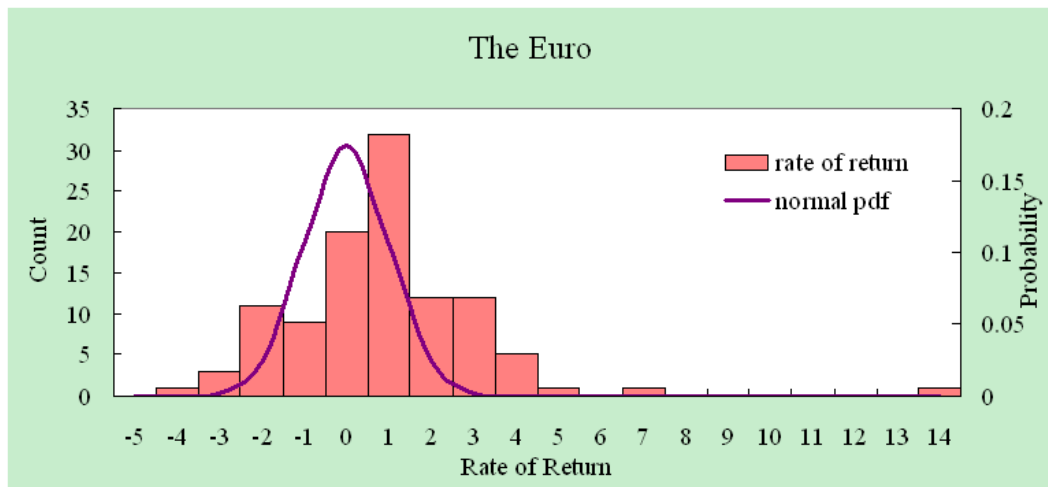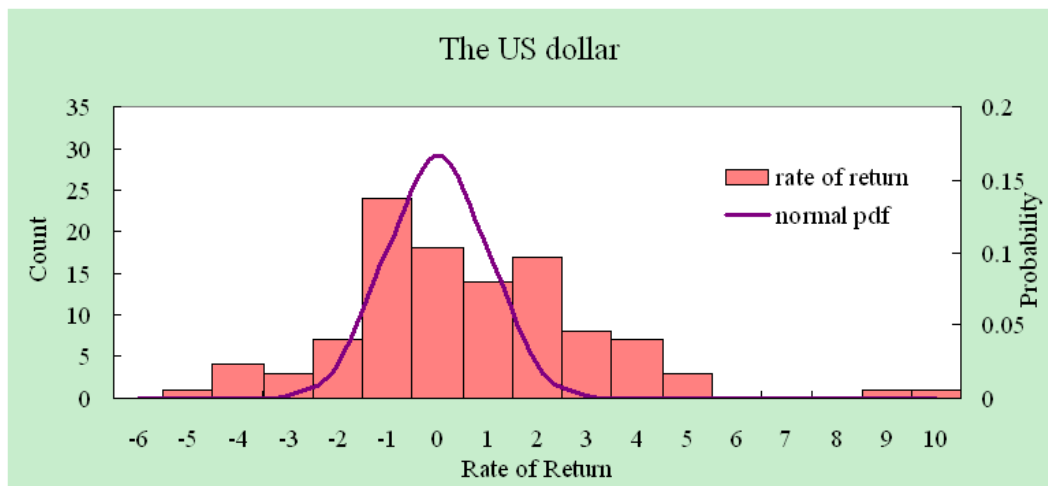
Figure 5.3

Figure 5.4

In general, if the histogram of the sample follows a bell-shaped pattern with the most of the frequency counts in the middle and the rest of the counts dying off out in the tails, then it is more likely that the underlying data can be modeled approximately by the

normal distribution. However, as shown in the Figure 5.3 and Figure 5.4, both histograms are not symmetric or in the shape of the bell. Therefore, it is not sufficient to assume the sample data of the foreign currency market follows the normal distribution, and a more sophisticated distribution is then required in order to model the behavior of this particular market.

5.1.2 Elliptical distribution

The theory for the elliptical distributions was first comprehensively investigated by Kelker [18] in 1970, later Muirhead [23] studied the elliptical distributions and also proposed some relevant robustness results of the elliptical distributions. Additional references in the field of the elliptical distributions are Tong [27] and Fang, Kotz and Ng [13].

The class of elliptical distributions is also recognized as the elliptical symmetric distributions or the elliptical contoured distributions, and this class contains distributions with similar features to the multivariate normal distribution but with either longer or shorter tails than the normal distribution. Several important distributions are known to belong to this family of distributions, such as the multivariate t, the multivariate Cauchy, the double exponential and the logistic distributions, and all of those distributions have elliptical contours of equiprobability.

Before the development of the so-called elliptical distributions, the multivariate normal distribution has been frequently used to model the multivariate data in different statistical areas for different purposes. Due to the reason that many properties of the multivariate normal distribution also extend to the elliptical distributions, hence, the class of the elliptical distribution becomes the alternative choice to describe the underlying multivariate data, especially in the case of the financial market.

This section will review some important and relevant information for the elliptical distributions that stated by Fang, Kotz and Ng [13] and also Tong [27].

Definition 5.5 *A random vector $y \in \mathbb{R}^n$ is spherical distributed if for every orthogonal*

matrix $O \in \mathbb{R}^{n \times n}$

$$Oy \overset{\mathrm{d}}{=} y$$

where the symbol $\ulcorner \overset{\mathrm{d}}{=} \lrcorner$ denotes the equality of the distributions.

In addition, the term $Y \sim S_n(\phi)$ is used to denote the random vector $y \in \mathbb{R}^n$ if the vector y is spherically distributed with the characteristic generator ϕ. On the other hand, if the random vector $y \in \mathbb{R}^n$ is spherically distributed with the density generator $g(y^T y)$, then the notation $Y \sim S_n(g)$ will be applied instead of $Y \sim S_n(\phi)$.

Definition 5.6 *A random vector $x \in \mathbb{R}^n$ with parameters $\mu \in \mathbb{R}^n$ and $\Sigma \in \mathbb{R}^{n \times n}$ is elliptically distributed if*

$$x \overset{\mathrm{d}}{=} \mu + A^T Y$$

where $Y \in \mathbb{R}^k$ is spherically distributed with the characteristic generator ϕ, $A \in \mathbb{R}^{k \times n}$ with $A^T A = \Sigma$, and $rank(\Sigma) = k$.

Furthermore, the corresponding notation that designates a random vector $x \in \mathbb{R}^n$ has the elliptical distribution with the characteristic generator ϕ and parameters $\mu \in \mathbb{R}^n$ and $\Sigma \in \mathbb{R}^{n \times n}$ is $X \sim EC_n(\mu, \Sigma, \phi)$.

Generally, an elliptically distributed random vector $x \in \mathbb{R}^n$ does not necessarily possess a probability density function. However, if there is a probability density function for the elliptically distributed random vector $x \in \mathbb{R}^n$, then the necessary condition $rank(\Sigma) = n$ must be fulfilled. And the density function of the elliptically distributed random vector $x \in \mathbb{R}^n$ is of the form

$$f_{\mu,\Sigma}(x) = |\Sigma|^{-\frac{1}{2}} g\left((x - \mu)^T \Sigma^{-1} (x - \mu) \right)$$

where $g : \mathbb{R} \to [0, \infty)$ is the nonincreasing density generator.

As mentioned earlier in this section, there are several distributions belong to the family of the elliptical distributions. And by considering the density generator g in the following form for $x \in \mathbb{R}$

$$g(x) = (2\pi)^{-\frac{n}{2}} \exp^{-\frac{1}{2}x},$$

the multivariate normal distribution can be proven to be one of the elliptical distributions as well.

Some useful results of the elliptical distributions will be summarized in the following remark.

Remark 5.7

- *If $x \sim EC_n(\mu, \Sigma, \phi)$ and $rank(\Sigma) = k$, then the elliptically distributed random vector $x \in \mathbb{R}^n$ has a stochastic representation*

$$x \stackrel{d}{=} \mu + rA^T \mu^{(k)}$$

with $A^T A = \Sigma$ and $r \geq 0$ being independent of $\mu^{(k)}$, where $\mu^{(k)}$ denotes a random vector that distributed uniformly on the unit sphere surface in $\mathbb{R}$.

- *If a distribution belongs to the family of the elliptical distributions, then all of its marginal and conditional distributions belong to this family, furthermore, any linear transformation of the elliptically distributed random vector x is again elliptically distributed. And all of those results for the elliptical distributions also apply to the multivariate normal distribution.*

- *If $x_i \sim EC_S(\mu, \Sigma, \phi)$ is independent and identically distributed for $i = 1, 2, \cdots, S$, then*

$$Z = \sum_{i=1}^{S} x_i \sim EC_S(S\mu, \Sigma, \phi^S)$$

with $\phi^S = \prod_{i=1}^{S} \phi$.

5.2 Estimations of the parameters

General speaking, once a statistical distribution has been selected for describing the underlying sample, the next approach is to estimate the parameters of the chosen distribution for the optimization problem.

The maximum likelihood estimator (MLE) is considered as one of the more robust parameter estimators, and hence widely applied in most of the practical applications. Except the maximum likelihood estimator, there are quite a few different choices of the parameter estimator. For example, Jorion [17] suggested to use the Bayesian estimators which combine the traditional parameter estimator together with the external prior information, and on the other hand, Jobson and Korkie [16] proposed that the more stable results can be achieved by using the Stein-type estimator.

This section will only consider and discuss the maximum likelihood estimator as the estimator for the mean and the covariance matrix, and the fundamental conception behind the maximum likelihood estimator is to determine the most likely values of the parameters that will best describe the sample data for a given distribution. That is, let x be a random variable with the probability density function

$$f(x; \theta_1, \theta_2, \cdots, \theta_n)$$

where $\theta_1, \theta_2, \cdots, \theta_n$ are the parameters that need to be estimated, and the maximum likelihood estimators of $\theta_1, \theta_2, \cdots, \theta_n$ are obtained by maximizing the likelihood function

$$L(\theta) = L(\theta_1, \theta_2, \cdots, \theta_n | x) = f(x; \theta_1, \theta_2, \cdots, \theta_n).$$

Furthermore, if the likelihood function is differentiable, then the maximum likelihood

estimator can be obtained by solving the maximum likelihood equation

$$\frac{d}{d\theta} \ln L(\theta) = 0.$$

Recall the definition of the normal distribution, a random variable X is normally distributed, $X \sim N(\mu, \sigma^2)$, if it has the probability density function

$$f(x; \mu, \sigma^2) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

with mean μ and variance σ^2. And by consider a set of random variables from the normal distribution based on a random sample of size n, $X_i \sim N(\mu, \sigma^2)$, the corresponding probability density function is

$$f(x_1, x_2, \cdots, x_n; \mu, \sigma^2) = \left(2\pi\sigma^2\right)^{-\frac{n}{2}} \exp\left(-\frac{\sum_{i=1}^{n}(x_i - \mu)^2}{2\sigma^2}\right)$$

and the likelihood function is

$$
\begin{aligned}
L(\mu, \sigma^2) &= f(x_1, x_2, \cdots, x_n; \mu, \sigma^2) \\
&= \left(2\pi\sigma^2\right)^{-\frac{n}{2}} \exp\left(-\frac{\sum_{i=1}^{n}(x_i - \mu)^2}{2\sigma^2}\right)
\end{aligned}
$$

In this distribution, there are only two parameters need to be considered, the mean μ and the variance σ^2, and the value of the maximum likelihood estimator for each parameters can be verified by solving the maximum likelihood equation with respect to the particular parameter. The maximum likelihood equation for the parameter μ is

$$
\begin{aligned}
0 &= \frac{d}{d\mu} \ln L(\mu, \sigma^2) \\
&= \frac{d}{d\mu} \ln \left((2\pi\sigma^2)^{-\frac{n}{2}} \exp\left(-\frac{\sum_{i=1}^{n}(x_i - \mu)^2}{2\sigma^2} \right) \right) \\
&= \frac{d}{d\mu} \left(\ln (2\pi\sigma^2)^{-\frac{n}{2}} - \frac{\sum_{i=1}^{n}(x_i - \mu)^2}{2\sigma^2} \right) \\
&= 0 - \frac{-2\sum_{i=1}^{n}(x_i - \mu)}{2\sigma^2} \\
&= \frac{2\sum_{i=1}^{n}(x_i - \mu)}{2\sigma^2}
\end{aligned}
$$

Hence, the maximum likelihood estimator for the parameter μ is

$$
\widehat{\mu}_{ML} = \overline{x} = \frac{\sum_{i=1}^{n} x_i}{n}.
$$

Similarly, the maximum likelihood estimator of the parameter σ^2 can be obtained by setting $\sigma^2 = \theta$.

$$
\begin{aligned}
0 &= \frac{d}{d\sigma^2} \ln L(\mu, \sigma^2) \\
&= \frac{d}{d\theta} \ln L(\mu, \theta) \\
&= \frac{d}{d\theta} \ln \left((2\pi\theta)^{-\frac{n}{2}} \exp\left(-\frac{\sum_{i=1}^{n}(x_i - \mu)^2}{2\theta} \right) \right) \\
&= \frac{d}{d\theta} \left(\ln (2\pi\theta)^{-\frac{n}{2}} - \frac{\sum_{i=1}^{n}(x_i - \mu)^2}{2\theta} \right) \\
&= \frac{d}{d\theta} \left(\text{constant} - \frac{n}{2} \ln\theta - \frac{\sum_{i=1}^{n}(x_i - \mu)^2}{2\theta} \right) \\
&= -\frac{n}{2\theta} + \frac{\sum_{i=1}^{n}(x_i - \mu)^2}{2\theta^2} \\
&= -\frac{n\theta}{2\theta^2} + \frac{\sum_{i=1}^{n}(x_i - \mu)^2}{2\theta^2}
\end{aligned}
$$

Equivalently,

$$
n\theta = \sum_{i=1}^{n}(x_i - \mu)^2
$$

Therefore, the maximum likelihood estimator for the parameter $\theta = \sigma^2$ is

$$\widehat{\theta}_{ML} = \widehat{\sigma^2}_{ML} = \frac{\sum_{i=1}^n (x_i - \widehat{\mu})^2}{n}.$$

For the elliptical distributed sample of size S with every random variables $x_i \in \mathbb{R}^n$, $i = 1, 2, \cdots, S$, being independent and identically distributed. The elliptical distributed sample can be written as

$$x_1 \sim EC_S(\mu, \Sigma, \phi)$$
$$x_2 \sim EC_S(\mu, \Sigma, \phi)$$
$$\vdots$$
$$x_S \sim EC_S(\mu, \Sigma, \phi)$$

with $x_1, x_2, \cdots, x_S \in \mathbb{R}^n$ and the characteristic generator ϕ assumed to be chosen as $-2\phi^T(0) = 1$. The corresponding maximum likelihood estimators for the mean μ and covariance Σ of this elliptically distributed sample are

$$
\begin{aligned}
\widehat{\mu}_{ML} &= \frac{\sum_{i=1}^S x_i}{S} \\
\widehat{\Sigma}_{ML} &= \frac{\sum_{i=1}^S (x_i - \widehat{\mu}_S)(x_i - \widehat{\mu}_S)^T}{S}
\end{aligned}
\tag{5.8}
$$

See Bain and Engelhardt [1] for more examples of the maximum likelihood estimators in different distributions, and Schöttle [26] for more information of different estimators in the elliptical distribution.

5.3   The example in the foreign currency market

The preceding sections have presented the preliminary requirements for defining the uncertainty set U of the portfolio optimization problem: the elliptical distributions for modeling the behavior of a financial market and the maximum likelihood estimator for determining the first two moments of the elliptically distributed random variable.

By considering the same foreign currency market example that has been discussed earlier in this paper, there are some notations need to be stated beforehand.

First of all, recall that the example of the foreign currency market consists only two foreign currencies, the Euro and the US dollar, and the historical data from January 2000 to December 2008 for the currencies will be adopted for calculating the corresponding parameters. This example uses the monthly rate of return of the exchange rate as the input sources instead of the exchange rates of the currencies.

Furthermore, the sample of the foreign currency market is assumed to be elliptical distributed with the Euro and the US dollar being independent and identically distributed. The vector $\mu = (\mu_e, \mu_{us}) \in \mathbb{R}^2$ denotes the expected monthly rate of returns and $\sigma^2 = (\sigma_e^2, \sigma_{us}^2)$ denotes the variance:

$$\mu = (\mu_e, \mu_{us})^T$$

$$= (0.368\%, 0.142\%)^T$$

$$\Sigma = \begin{pmatrix} \sigma_e^2 & \sigma_{eus} \\ \sigma_{eus} & \sigma_{us}^2 \end{pmatrix}$$

$$= \begin{pmatrix} 5.189\% & 1.138\% \\ 1.138\% & 5.714\% \end{pmatrix}$$

By denoting the vector $x = (x_e, x_{us})^T$ as the weights of the portfolio with $x_e + x_{us} = 1$. The expected return of the portfolio E and the variance of return on the portfolio V are

$$E = x^T \mu = x_e \mu_e + x_{us} \mu_{us} \tag{5.9}$$

$$V = x^T \Sigma x \tag{5.10}$$

And by substituting the value of the mean μ into equation (5.9) and the value of covariance matrix Σ into equation (5.10), the expected return of the portfolio E and the variance of

return on the portfolio V are given as

$$
\begin{aligned}
E &= x_e \mu_e + x_{us} \mu_{us} \\
&= 0.368 x_e \% + 0.142 x_{us} \% \\
V &= x^T \Sigma x \\
&= \begin{pmatrix} x_e & x_{us} \end{pmatrix} \begin{pmatrix} 5.189\% & 1.138\% \\ 1.138\% & 5.714\% \end{pmatrix} \begin{pmatrix} x_e \\ x_{us} \end{pmatrix} \\
&= 5.189 x_e^2 \% + 5.714 x_{us}^2 \% + 2.276 x_e x_{us} \%
\end{aligned}
$$

Secondly, the formulation for determining the optimal portfolio of the considered optimization problem needs to be defined. In the case of the foreign currency market, the risk aversion formulation (2.19) will be adopted throughout the rest of the dissertation, and the reason to chose the risk aversion formulation as the objective function for the optimization problem is that the structure of the risk aversion formulation explicitly model the trade-off between the expected return and risk by using the risk aversion coefficient ϱ.

Definition 5.11 *By assuming the shortselling is not allowed for this optimization problem, the original portfolio optimization problem (the original (GCP_u) optimization problem) can be defined as*

$$
\max_{x \in X} \ (x^T \mu - \lambda x^T \Sigma x)
$$

with

- *$X \subset \{x \in \mathbb{R}^2 \mid x^T = 1, \ x \geqslant 0\}$ be non-empty, convex and compact,*

- *$\mu \in \mathbb{R}^2$ denotes the expected return,*

- *$\Sigma \in \mathbb{R}^{2 \times 2}$ denotes the covariance matrix with $\Sigma \succ 0$,*

- $\varrho \geqslant 0$ *denotes the risk aversion coefficient.*

Remark 5.12 *Note that, the original (GCP_u) optimization problem reduces to the expected return maximizing optimization problem (2.18) as the risk aversion coefficient $\varrho = 0$, and conversely, when the coefficient ϱ is large, the original (GCP_u) optimization problem becomes to concern mostly the risk of the investment rather than the expected return.*

Strictly speaking, the mean μ and the covariance matrix Σ are the only possible uncertain parameters for the mean variance optimization problem. Therefore, the pair (μ, Σ) is supposed to be the uncertain parameter u for the original (GCP_u) optimization problem. However, the uncertainty set is usually defined only for the mean μ in most of the practical cases, and the reason of that is the covariance matrix Σ behaves as a less volatile and also less effective parameter for the mean variance optimization problem compared to the mean μ.

Definition 5.13 *The formulation for the local robust counterpart of the original (GCP_u) optimization problem is*

$$\max_{x \in X} \min_{v \in U_\delta(\widehat{\mu})} \ (x^T v - \lambda x^T \Sigma x)$$

with

- $X \subset \{x \in \mathbb{R}^2 \mid x^T = 1, \ x \geqslant 0\}$ *be non-empty, convex and compact,*

- $U_\delta(\widehat{\mu})$ *denotes an uncertainty set centered at point $\widehat{\mu}$ with the ratio of size $\delta \geqslant 0$,*

- $\varrho \geqslant 0$ *denotes the risk aversion coefficient.*

Next, to complete the formulation for the local robust counterpart of the original (GCP_u) optimization problem, an explicit description of the uncertainty set has to be made. As mentioned before, the form and structure of the uncertainty set U can be varied

depend on different requirements or concerns on future values of certain parameters. For instance, Tütüncü and Koenig [28] preferred the interval uncertainty set that obtained by using the bootstrapping strategies and also the moving averages of returns. On the other hand, Ben-Tal and Nemirovski [3] and Schöttle [26] proposed to solve the optimization problem with the ellipsoidal uncertainty set. And Goldfarb and Iyengar [15] considered a linear factor model for the multivariate returns of assets and defined the uncertainty set by using both interval and ellipsoidal uncertainty sets. However, despite all those possible options for defining the uncertainty set, the ellipsoidal uncertainty set is more intuitive to apply in the optimization problem by considering a respective point estimate as the center of the uncertainty set with the related covariance matrix and the desired level of confidence that denote the shape and the size respectively.

For the sample of the foreign currency market, the uncertainty set $U_\delta(\widehat{\mu})$ can be constructed by choosing the maximum likelihood estimator for the mean μ, $\widehat{\mu}_{ML}$, as the center of the ellipsoidal uncertainty set with the covariance matrix Σ and the ratio of size $\delta \geqslant 0$.

To be more specified, the maximum likelihood estimator $\widehat{\mu}_{ML}$ is elliptically distributed based on the independent and identically distributed random vectors $\mu = (\mu_e, \mu_{us})$. That is, by equation (5.8) the maximum likelihood estimator $\widehat{\mu}_{ML}$ is

$$\widehat{\mu}_{ML} \;\; = \;\; \frac{\mu_e + \mu_{us}}{2}$$

and according to remark 5.7, the distribution of the maximum likelihood estimator $\widehat{\mu}_{ML}$ is

$$\widehat{\mu}_{ML} = \frac{1}{2}(\mu_e + \mu_{us}) \sim EC_2(2\mu, \Sigma, \phi^2)$$

Equivalently,

$$\widehat{\mu}_{ML} \sim EC_2(\mu, \frac{1}{2}\Sigma, \phi^2) \tag{5.14}$$

By applying the definition of an ellipsoidal set[1], the uncertainty set $U(\widehat{\mu})$ can be displayed as

$$U(\widehat{\mu}) = \{v \in \mathbb{R}^2 \mid (v - E[\widehat{\mu}_{ML}])^T (Cov[\widehat{\mu}_{ML}])^{-1} (v - E[\widehat{\mu}_{ML}]) \leqslant \delta^2\}$$

and by substituting (5.14), the uncertainty set $U(\widehat{\mu})$ is in the form of

$$U(\widehat{\mu}) = \{v \in \mathbb{R}^2 \mid 2(v - \widehat{\mu})^T \widehat{\Sigma}^{-1} (v - \widehat{\mu}) \leqslant \delta^2\} \tag{5.15}$$

with δ^2 denotes the desired level of confidence, $\widehat{\mu} = (0.368\%, 0.142\%)$ and $\widehat{\Sigma} = \begin{pmatrix} 5.189\% & 1.138\% \\ 1.138\% & 5.714\% \end{pmatrix}$.

Alternatively, the uncertainty set $U(\widehat{\mu})$ can also be expressed as

$$U(\widehat{\mu}) = \{v \in \mathbb{R}^2 \mid v = \widehat{\mu} + \frac{\delta}{\sqrt{2}} \Sigma^{\frac{1}{2}} w, \ \|w\| \leqslant 1\} \tag{5.16}$$

by introducing a vector $w \in \mathbb{R}^2$ with

$$w = \frac{\sqrt{2}}{\delta} \Sigma^{-\frac{1}{2}} (v - \widehat{\mu}) \ \text{ and } \ w^T w = \|w\|^2 = 1.$$

Note that, the formulation for the local robust counterpart of the original (GCP_u) optimization problem can be rearranged by applying the uncertainty set $U(\widehat{\mu})_\delta$ stated in (5.16). That is,

$$\max_{x \in X} \min_{v \in U_\delta(\widehat{\mu})} \left(x^T v - \lambda x^T \Sigma x \right)$$

---

[1] An ellipsoidal set in $\mathbb{R}^n$ is defined as

$$\varepsilon = \{x \in \mathbb{R}^n \mid (x - x_c)^T A^{-1} (x - x_c) \leqslant 1\}$$

The vector $x_c \in \mathbb{R}^n$ denotes the center of the ellipsoid and the matrix $A = A^T \succ 0$ denotes the distance that the ellipsoid extends in every direction from x_c.

by substituting $v = \widehat{\mu} + \frac{\delta}{\sqrt{2}} \Sigma^{\frac{1}{2}} w$

$$
\begin{aligned}
&= \max_{x \in X} \min_{\|w\| \leqslant 1} \left(x^T \left(\widehat{\mu} + \frac{\delta}{\sqrt{2}} \Sigma^{\frac{1}{2}} w \right) - \lambda x^T \Sigma x \right) \\
&= \max_{x \in X} \min_{\|w\| \leqslant 1} \left(x^T \widehat{\mu} + \frac{\delta}{\sqrt{2}} x^T \Sigma^{\frac{1}{2}} w - \lambda x^T \Sigma x \right) \\
&= \max_{x \in X} \left(x^T \widehat{\mu} - \lambda x^T \Sigma x + \min_{\|w\| \leqslant 1} \frac{\delta}{\sqrt{2}} x^T \Sigma^{\frac{1}{2}} w \right)
\end{aligned}
$$

Since the last term of equation $\min_{\|w\| \leqslant 1} \frac{\delta}{\sqrt{2}} x^T \Sigma^{\frac{1}{2}} w$ is minimized if $w = -\frac{\Sigma^{\frac{1}{2}} x}{\|\Sigma^{\frac{1}{2}} x\|}$. Hence,

$$
\begin{aligned}
&= \max_{x \in X} \left(x^T \widehat{\mu} - \lambda x^T \Sigma x - \frac{\delta}{\sqrt{2}} x^T \Sigma^{\frac{1}{2}} \frac{\Sigma^{\frac{1}{2}} x}{\|\Sigma^{\frac{1}{2}} x\|} \right) \\
&= \max_{x \in X} \left(x^T \widehat{\mu} - \lambda x^T \Sigma x - \frac{\delta}{\sqrt{2}} \|\Sigma^{\frac{1}{2}} x\| \right) \\
&= \max_{x \in X} \left(x^T \widehat{\mu} - \lambda x^T \Sigma x - \frac{\delta}{\sqrt{2}} \sqrt{x^T \widehat{\Sigma} x} \right) \quad\quad\quad (5.17)
\end{aligned}
$$

Finally, after reformulating the local robust counterpart of the original (GCP_u) optimization problem by using the uncertainty set $U(\widehat{\mu})_\delta$, the efficient portfolios for the sample of the foreign currency market can be then obtained by choosing the desired values of δ and ϱ.

To start with, the efficient frontier of the original (GCP_u) optimization problem will be obtained in order to compare with the results of the robustification. And since for each different value of ϱ, an efficient portfolio is generated by solving the original (GCP_u) optimization problem. Therefore, the entire efficient frontier of the original (GCP_u) optimization problem can be obtained by increasing ϱ from zero.

Figure 5.18 presents the possible asset allocations of the investment in the foreign currencies. The original (GCP_u) efficient frontier starts from the global minimum variance portfolio (GMV) with investment proportion 53% in Euros and 47% in US dollars, and the end point of the efficient frontier is the maximum return portfolio (MRP) with 100% in Euros. In addition, the curve below the (GMV) are the inefficient portfolios.

Figure 5.18
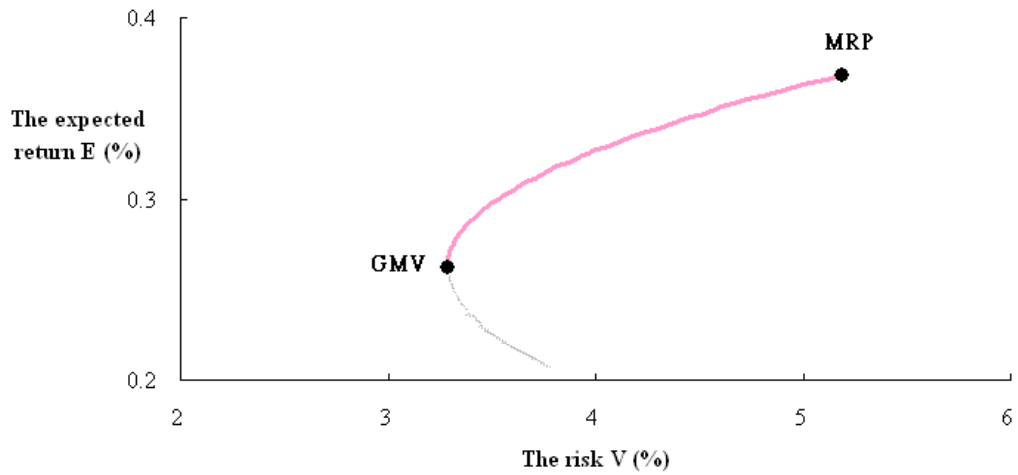
For the robustification of the portfolio optimization problem, the value for the level of confidence(ratio of size) $\delta \geqslant 0$ will be chosen between zero and one. The following figure shows the robust efficient frontier for $\delta = 0.2$ together with the original (GMV) efficient frontier.

Figure 5.19
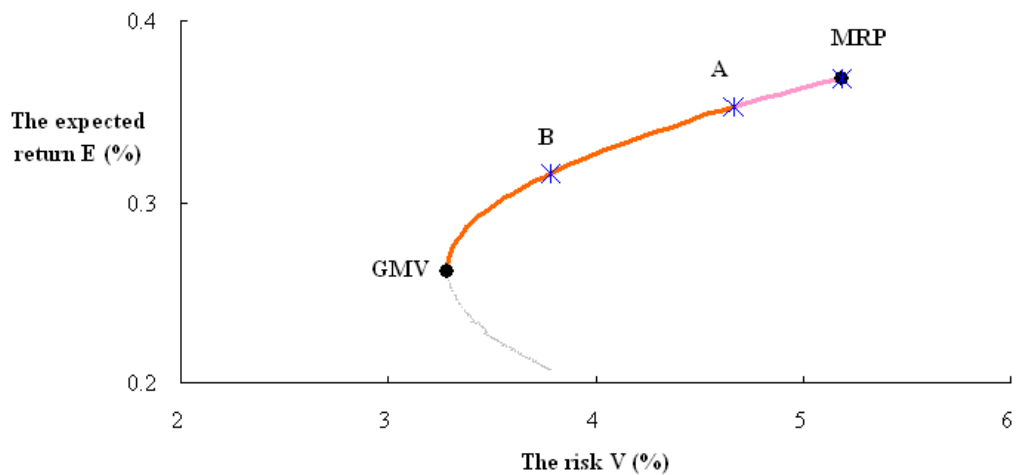
It can be easily observed from figure 5.19 that the pattern of the efficient frontier for the robustified optimization problem is equal to the original (GCP_u) efficient frontier, but the length of the efficient frontier for the robust optimization problem varies inversely as the level of confidence δ increases. For instance, the robust efficient frontier is exactly

the same with the original (GCP_u) efficient frontier if $\delta = 0.1$. On the other hand, the robust efficient frontier with $\delta = 0.2$ begins from the point (GMV) to point A, and for $\delta = 0.3$, the robust efficient frontier starts from the point (GMV) and finished at point B. However, except the graphical meaning that the set of efficient portfolios decreases as the level of confidence δ increases. This phenomenon also shows that the obtained robust optimal portfolio moves towards to the global minimum variance portfolio as the size of the uncertainty set increases. As shown in the figure 5.20 below, for a fixed risk aversion coefficient $\lambda = 0.05$, the robust optimal portfolios A(with $\delta = 0.2$) and B(with $\delta = 0.5$) are closer to the (GMV) portfolio than the original (GCP_u) optimal portfolio.
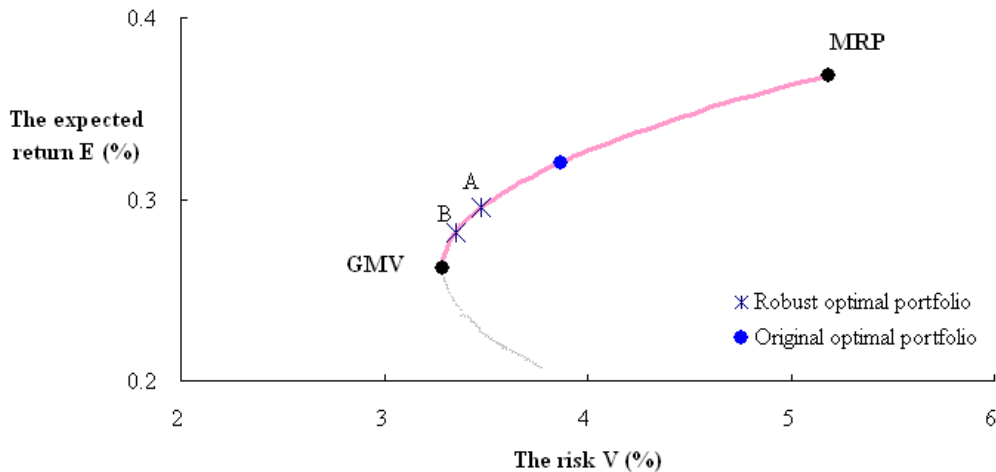
Figure 5.20

Figure 5.21 is displayed for the purpose of showing the comparison between the optimal portfolio allocations with the risk aversion coefficient $\varrho = 0.05$. As mentioned earlier, the robust optimal portfolio obtained by choosing a more secure uncertainty set $U(\widehat{\mu})_\delta$ will be more conservative, consequently, the expected return of the portfolio will be lower. This fact can be also observed from the figure below, the investment in the US dollar (the asset with lower variance) becomes progressively heavier as the level of confidence δ increases. In other words, the process of applying the local robust counterpart approach to the portfolio optimization problem can be ended up as finding the (GMV) portfolio of the original (GCP_u) optimization problem if the level of the confidence δ is too high.

Figure 5.21 *The optimal portfolio allocation with $\varrho = 0.05$*
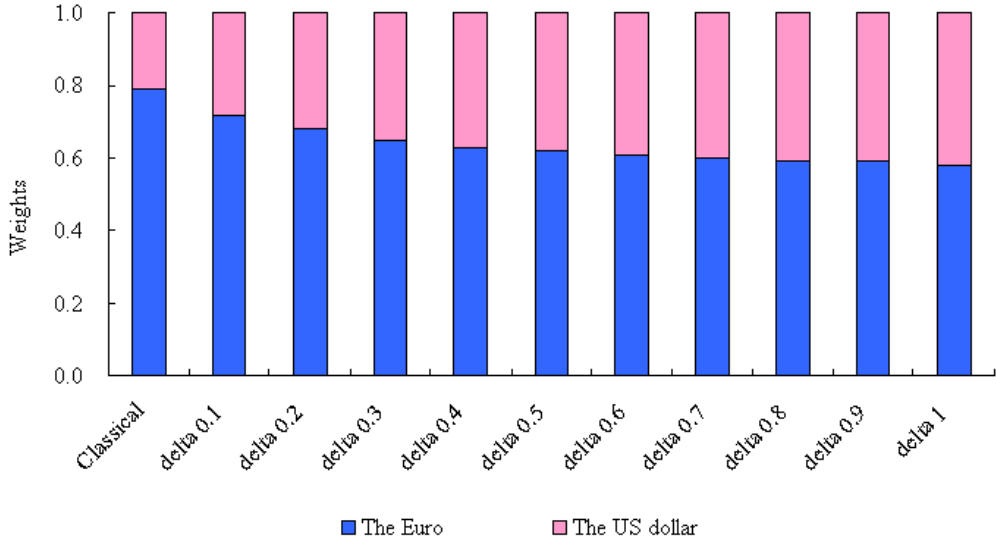
CHAPTER 6

CONCLUSION

The main motivation of this dissertation is to apply the local robust counterpart approach to the famous Markowitz mean variance portfolio optimization problem. To enable this application, the concepts of the mean variance portfolio optimization, the general convex conic optimization, and also the (local) robust counterpart approach have been studied and discussed. As the purpose of this dissertation is to investigate the process of applying the robust counterpart approach, therefore, only the related and also necessary notions and properties were mentioned.

The robustification of an optimization problem provides more secured and stable results even if the underlying parameters are uncertain. However, as observed from the given example of the foreign currency market, the obtained optimal portfolio can be too conservative if the chosen uncertainty set is too large, and on the other hand, if the uncertainty set is too small, then the robust counterpart approach will not be sufficiently to apply for optimizing the problem.

Above all, in view of the fact that the (local) robust counterpart approach is the methodology that optimizes the problem over the uncertainty set, therefore, explicitly definitions and restrictions are still needed in order to characterize the uncertainty set without eliminate the important information from the original data. Although there are some recommended guidelines for constructing the uncertainty set for the robust optimization problem, the investigation of how to creat an appropriate uncertainty set for

the particular optimization problem is still in progress.

Appendix A

The following tables provide the historical data of the Euro and the US dollar from January 2000 to December 2008.

The monthly rate of return of the Euro (%)									
	2000	**2001**	**2002**	**2003**	**2004**	**2005**	**2006**	**2007**	**2008**
Jan	-2.959	1.434	-2.560	0.601	-3.013	-2.041	-0.828	-1.352	0.094
Feb	1.212	0.236	0.443	4.505	-2.047	-0.361	0.103	1.982	1.341
Mar	-1.752	-2.769	0.393	0.612	0.436	0.116	2.512	0.937	3.978
Apr	-2.803	-0.209	1.095	0.736	1.502	-1.501	-0.645	0.917	-0.405
May	6.471	-4.148	3.034	2.661	-1.114	-0.602	-0.131	0.015	0.511
Jun	1.491	0.719	0.466	-3.024	0.857	0.000	0.758	-0.737	0.126
Jul	-1.543	2.609	-2.819	0.358	-1.165	2.146	-1.541	-0.458	-0.442
Aug	-0.940	2.020	0.444	-0.829	2.519	-0.870	-1.363	0.297	2.397
Sep	-3.068	-0.705	-0.977	0.387	1.149	0.000	0.788	3.514	-2.811
Oct	-2.750	0.016	0.909	-2.504	0.609	-0.295	-1.021	-0.100	-0.797
Nov	3.629	1.323	0.377	2.140	0.014	0.044	0.612	2.956	2.734
Dec	2.954	-2.425	1.971	1.382	2.550	1.507	0.074	3.554	13.912

The monthly rate of return of the US dollar (%)									
	2000	**2001**	**2002**	**2003**	**2004**	**2005**	**2006**	**2007**	**2008**
Jan	0.883	2.333	2.329	-2.796	-1.735	-0.132	-2.157	0.789	-0.435
Feb	2.244	2.468	-0.028	4.065	-1.965	-2.369	1.619	0.532	-1.043
Mar	-0.666	2.328	-0.397	0.000	1.701	1.555	1.069	-0.176	-0.258
Apr	2.042	-0.724	-1.138	-1.464	4.756	-1.096	-4.430	-1.088	-0.237
May	4.483	0.874	-0.088	-1.643	-2.118	3.744	-1.711	1.060	-0.079
Jun	-1.403	0.580	-4.614	-1.013	1.323	1.289	1.009	-1.366	-1.316
Jul	1.075	-0.694	-2.080	3.425	-0.200	0.638	-1.217	-0.906	0.518
Aug	3.120	-1.289	0.638	1.672	1.516	-1.593	-1.793	0.609	8.231
Sep	-1.183	-1.234	-1.426	-5.421	-0.861	3.421	1.884	-0.949	0.935
Oct	1.175	1.671	0.330	-1.734	-1.994	-0.860	-1.169	-1.655	9.579
Nov	2.070	2.559	0.548	-2.236	-4.125	1.797	-3.051	1.249	3.281
Dec	-4.197	-2.373	-3.375	-3.840	0.444	0.761	1.208	3.845	3.220

The average monthly rate of return of the Euro and the US dollar (%)									
	2000	**2001**	**2002**	**2003**	**2004**	**2005**	**2006**	**2007**	**2008**
EUR	-0.005	-0.158	0.231	0.585	0.191	-0.155	-0.057	0.96	1.72
USD	0.804	0.542	-0.775	-0.915	-0.271	0.596	-0.728	0.162	1.866

Appendix B

For a given pair $(\widehat{u}, \delta)$ with $u \in U_\delta(\widehat{u})$ and $\delta \geqslant 0$, the associated sets of the $(LRC_{\widehat{u},\delta})$ problem are:

- The feasible set $F_{LRC}(\widehat{u}, \delta) : \{x \in X \mid g(x, u) \preceq_K 0\}$.

- The optimal set $F_{LRC}^*(\widehat{u}, \delta) : \{x \in F_{LRC}(\widehat{u}, \delta) \mid f_{LRC}(x, \widehat{u}, \delta) \leq f_{LRC}^*(\widehat{u}, \delta)\}$.

- The set of Slater points $F_{LRC}^S(\widehat{u}, \delta) : \{x \in X \mid g(x, u) \prec_K 0\}$.

where $f_{LRC}^* : U \times \mathbb{R}_+ \to \mathbb{R}$ is the optimal value function defined as

$$f_{LRC}^*(\widehat{u}, \delta) := \min\{f_{LRC}(x, \widehat{u}, \delta) \mid x \in F_{LRC}(\widehat{u}, \delta)\}. \tag{6.1}$$

Furthermore, the corresponding functions assign the feasible set $F_{LRC}(\widehat{u}, \delta)$ and the optimal set $F_{LRC}^*(\widehat{u}, \delta)$ are

- The feasible set mapping $F_{LRC} : U \times \mathbb{R}_+ \to \wp(X)$.

- The optimal set mapping $F_{LRC}^* : U \times \mathbb{R}_+ \to \wp(X)$.

List of References

[1] L.J. Bain and M. Engelhardt. *Introduction to Probability and Mathematical Statistics.* Duxbury Press, California, second edition, 2000.

[2] B. Bank. *Non-linear Parametric Optimization.* Birkhauser, Basel, 1983.

[3] A. Ben-Tal and A. Nemirovski. Robust convex optimization. *Mathematics of Operations Research*, 23(4):769–805, 1998.

[4] A. Ben-Tal and A. Nemirovski. Selected topics in robust convex optimization. *Mathematical Programming*, 112(1):125–158, 2008.

[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. *Review of Financial Studies*, 4(2):315–342, 1991.

[6] S. Boyd and L. Vandenberghe. *Convex Optimization.* Cambridge University Press, 2004.

[7] J. Brinkhuis and V. Tikhomirov. *Optimization: Insights and Applications.* Princeton University Press, New Jersey, 2005.

[8] M. Britten-Jones. The sampling error in estimates of mean-variance efficient portfolio weights. *Journal of Finance*, 54(2):655–671, 1999.

[9] G. Calafiore and M.C.Campi. Uncertain convex programs: randomized solutions and confidence levels. *Mathematical Programming*, 102(4):25–46, 2005.

[10] G. Cornuejols and R. Tutuncu. *Optimization Methods in Finance.* Cambridge University Press, Cambridge, 2007.

[11] E.J. Elton, M.J. Gruber, S.J. Brown, and W.N. Goetzmann. *Modern Portfolio Theory and Investment Analysis.* John Wiley & Sons, New York, sixth edition, 2003.

[12] F.J. Fabozzi, P.N. Kolm, D.A. Pachamanova, and S.M. Focardi. *Robust Portfolio Optimization and Management.* John Wiley & Sons, New Jersey, 2007.

[13] K.-T. Fang, S. Kotz, and K.-W. Ng. *Symmetric Multivariate and Related Distributions.* Chapman and Hall, London, 1990.

[14] G.M. Frankfurter, H.E. Phillips, and J.P. Seagle. Portfolio selection: The effects of uncertain means, variances and covariances. *Journal of Financial and Quantitative Analysis*, 6:1251–1262, 1971.

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

[16] J.D. Jobson and B. Korkie. Putting Markowitz theory to work. *Journal of Portfolio Management*, 7(4):70–74, 1981.

[17] P. Jorion. Bayes-Stein Estimation for Portfolio Analysis. *Journal of Financial and Quantitative Analysis*, 21(3):279–292, 1986.

[18] D. Kelker. Distribution theory of spherical distributions and a location-scale parameter. *Sankhya : The Indian Journal of Statistics*, 32(A):499–516, 1970.

[19] L. Laloux, P. Cizeau, J.P. Bouchaud, and M. Potters. Noise dressing of financial correlation matrices. *Physical Review Letters*, 83:1467–1470, 1999.

[20] H.M. Markowitz. *Portfolio Selection: Efficient Diversification of investments*. John Wiley & Sons, New York, 1959.

[21] H.M. Markowitz. *Mean-Variance Analysis in Portfolio Choice and Capital Markets*. Blackwell, New York, 1987.

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

[23] R.J. Muirhead. *Aspects of Multivariate Statistical Theory*. John Wiley, New York, 1982.

[24] J.M. Mulvey, R.J. Vanderbei, and S.A. Zenios. Robust optimization of large-scale systems. *Operations Research*, 43(2):264–281, 1995.

[25] S. Pafka and I. Kondor. Estimated correlation matrices and portfolio optimization. *Physica A*, 343:623–634, 2004.

[26] K. Schöttle. *Robust Optimization with Application in Asset Management*. PhD thesis, Technische Universität München, Department of Mathematics, 2007.

[27] Y.L. Tong. *The Multivariate Normal Distribution*. Springer-Verlag, New York, 1990.

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