千葉工業大学

博士学位論文

Portfolio Selection based on Cumulative Prospect Theory

(累積プロスペクト理論に基づくポートフォリオ選択)

平成 29 年 9 月

キョウ　チョウ

Abstract

Cumulative prospect theory (CPT) has become one of the most popular approaches for evaluating the behavior of decision makers under conditions of uncertainty. Substantial experimental evidence suggests that human behavior may significantly deviate from the traditional expected utility maximization framework when faced with uncertainty. The problem of portfolio selection should be therefore revised when the investor's preference is for CPT instead of expected utility theory.

CPT can describe the behavior of bounded rational decision makers in a psychologically more realistic way, over the past decade, researchers in the field of behavioral economics have repeatedly considered how CPT should be applied in economic settings; these efforts are now bearing fruit. Although CPT has received a great deal of attention, to the best of our knowledge, little research has investigated the portfolio choice problem based on CPT due to the complexity of CPT function.

The purposes of this dissertation evaluated the problem of identifying the optimal portfolio consisting of one riskless asset and multiple risky assets under CPT. The CPT function is generally non-convex, non-concave and non-smooth, which means that traditional optimization methods such as Lagrange multipliers and convex duality do not work and the CPT function may have many local maxima. A real-coded genetic algorithm was to be used to solve the problem of portfolio choice. To overcome the limitations of RCGA and improve its performance, an adaptive method and a new selection operator were introduced. Computational results show that the new method is a rapid, effective, and stable genetic algorithm with the influence of various parameters on the CPT values being presented.

A method which couples scenario techniques for simulating the scenario of the real stock market with a genetic algorithm to determine the optimal solution was presented. The major challenge is to provide data on mathematical models in determining optimal solutions to address uncertainties in the field of financial investment. The effectiveness of the mathematical models hinges on the quality of the scenarios. This dissertation focused on three different variants of the bootstrap method for scenario generation. Bootstrap method being a form of resampling in statistics, it is a highly effective tool in the absence of a parametric distribution for a set of data and suitable for assessing the distribution properties of some statistic of such data.

Financial regulators now propose some risk management requirements in terms of loss. Mathematically，risk management is a process of how to control the loss distributions. The value-at-risk (VaR) and conditional value-at-risk (CVaR) are popular tools for managing risk. This dissertation analyzed the portfolio optimization under CPT with risk constraints, deviation constraints, and other constraints. And, the optimal portfolios under various constraints were given in this dissertation. It found that CPT investor with constraints of risk and deviation significantly changed their investment behavior. Moreover, due to the constraints of risk and deviation, the CPT value decreased and the investment income declined.

Acknowledgements

Firstly, I would like to express my sincere gratitude to my advisor Prof. Chunhui Xu for the continuous support of my Ph.D study and related research, for his patience, motivation, and immense knowledge. His guidance helped me in all the time of research and writing of this thesis. I could not have imagined having a better advisor and mentor for my Ph.D study.

Besides my advisor, I would like to thank the rest of my thesis committee: Prof. Akiya Inoue, Prof. Iwashita Motoi, Prof. Shirai Yutaka, and Prof. Tsutomu Kounosu, for their insightful comments and encouragement, but also for the hard question which incented me to widen my research from various perspectives.

My sincere thanks also go to Prof. Shuning Wang, Prof. Masakazu Ando, Ji Wang, Dr. Xiangming Xi, Dr. Hao Chen, and Shuren Bi, who provided me a few suggestions. Without their precious support it would not be possible to conduct this research well.

A very special gratitude goes out to Mitsui Bussan Trade Promotion Foundation for offering scholarships.

Last but not the least, I would like to thank my family: my parents and to my wife and daughter for supporting me spiritually throughout writing this thesis and my life in general.

Contents

List of Figures

List of Tables

List of Abbreviations

ADF	Augmented Dickey–Fuller
ARCGA	Adaptive Real-Coded Genetic Algorithm
BCGA	Binary-Coded Genetic Algorithm
CAL	Capital Allocation Line
CPT	Cumulative Prospect Theory
CV	CPT Value
CVaR	Conditional Value-at-Risk
DTNS	Duplicated Top-N Selection
EUT	Expected Utility Theory
EV	Expected Value
FR	Final Return
GA	Genetic Algorithm
JB	Jarque-Bera
MBB	Moving Block Bootstrap
MPT	Modern Portfolio Theory
MV	Mean-Variance
NBB	Non-Overlapping Block Bootstrap
PT	Prospect Theory
Q-Q	Quantile–Quantile
RAM	Random Access Memory
RCGA	Real-Coded Genetic Algorithm
RRWS	Rank-based Roulette Wheel Selection
SB	Standard Bootstrap
VaR	Value-at-Risk
WTP	Willingness to Pay

Chapter 1

Introduction

1.1 Portfolio Selection

In finance, portfolio choice is a process of allocating one's investable wealth to various financial assets according to some optimality criteria determining the best possible trade-off between the return (or utility) and certain constraint conditions，such as resources or risk. Further, there are two main objectives: one question is that how to determine a proportion to invest in each type of asset within the portfolio for receiving the highest possible return; the other one is that appropriate level of risk should be considered for given return.

Modern portfolio theory (MPT), or mean-variance analysis, has been proposed by Markowitz (1952) to provide the theoretical background for the relationship between the risk and return of a portfolio. Also the theory provides a mathematical framework assuming that investors are risk averse and make choice in terms of the expected return and its variance, giving important insight that an asset's return and risk should not be assessed by itself, and contributing to a portfolio's overall return and risk. An efficient frontier curve is constructed by varying the weights for each asset and recalculating the expected and standard deviation. MPT has a high value for portfolio management, because rational investors will always choose to invest on this frontier according to trade-off between return and risk, i.e., risk attitude. MPT has been thought of as the beginning of modern financial economics, because there was no concept of investment portfolio before the 1950s (Rubinstein, 2002).

Sharpe ratio, one of the most famous concepts in finance, was proposed by Sharpe (1966) based on the further expansion of MPT. Under MPT, people has realized that portfolio problems should be taken into account the efficient frontier. According to their risk attitudes, people can select one point on the efficient frontier according to tradeoff between risk and return. As a risk-free asset is introduced to the mix, there is just one portfolio of risky assets to be held by people. Capital allocation line (CAL) combining the risk-free asset with the risky assets should therefore be fully considered when people make choice.

Although the fact of maximizing expected return has never been disputed, it is difficult to have unified standard for judging the risk. Thus different people have different understanding about the it. Markowitz used the standard deviation of return as a risk measure. However, the using standard variance, or variance, has some drawbacks. Thus it aroused some arguments about standard variance as a good indicator for measuring risk and many new methods for measuring risk have been developed. Value-at-risk (VaR) as one of the most popular tools has emerged in 1994(JPMorgan, 1994, 1997). Jorion (2000), Linsmeier and Pearson (2000), Alexander and Baptista (2002),Chance (2004), and Hull (2008) noticed that VaR has become a popular risk management tool by corporate treasurers, dealers, fund managers, financial institutions, and regulators such as Basle Committee on Banking Supervision. Alexander and Baptista (2002) related VaR to mean-variance analysis and examined the economic implications of using a mean-VaR model for portfolio selection. They found that mean-variance efficient portfolios with the higher variance portfolio might have less VaR. Mean-VaR model, or mean-VaR efficient frontier, was accepted as a kind of popular tools for maximizing profits at given a specific VaR level.

However, the VaR as a measure for risk is under debate. Rockafellar and Uryasev (2000) introduced a new method, called Conditional value-at-risk (CVaR), to manage risk. Although CVaR is similar to VaR risk measurement, they are based on different mathematical properties. Artzner et al. (1999) showed that VaR is not a "coherent" measure of risk because it fails to satisfy the "subadditivity property." Rockafellar and Uryasev (2000) and Rockafellar and Uryasev (2002) showed that CVaR is superior to VaR in optimization applications. For these reasons, some researchers have thought that CVaR should be used rather than VaR as a tool of measuring risk. However, the other scholars have proposed that there are advantages and disadvantages to both approaches, even though conclusions made from both of them may be contradicted. (Alexander and Baptista, 2004, Sarykalin et al., 2008).

Beyond MPT, expected utility theory (EUT) is a widely accepted as a well-known theory for explaining portfolio choice when faced with uncertainty. In fact, EUT had a long history since 18th century and was originated from gambling in casinos. At that time, people thought that rational gamblers should be based on the expected return of outcomes, i.e. expected value theory. However, this theory has been questioned by the famous St. Petersburg paradox. Daniel Bernoulli gave an explanation that people do not make decisions based on earnings, but on the moral expectations of income. As a result, this explanation is the prototype of expected utility theory (Bernoulli, 1954).

EUT was formally developed by Neumann et al. (1944) in their book "Theory of Games and Economic Behavior", saying the main concern of EUT is the representation of individual attitudes towards risk and it has been a predominant model for portfolio choice based on the assumption that people are rational (Karni, 2014). Under EUT, people start to evaluate wealth according to final asset positions and treat probability objectively. Furthermore, people are of uniformly risk aversion. There are many papers discussing the portfolio optimization under EUT, such as Merton (1969), Samuelson (1969), Duffie (2010),Karatzas et al. (1998), Merton and Samuelson (1990), Föllmer and Schied (2011).

These theories have been very useful in modeling portfolio choice and substantial empirical and experimental evidence (such as the paradoxes outlined by Allais (1953) and Ellsberg (1961) has revealed that they do not reflect reality. Because of the assumption that people are rational, these theories merely demonstrate how people should behave instead of their actual portfolio choices under risk. Prospect theory (PT) proposed by Kahneman and Tversky (1979) is still widely viewed as the best available description of people's actual behavior when evaluating risk in experimental settings, particularly when psychological insights are incorporated (Barberis, 2013).

Inspired by Quiggin (1982), Tversky and Kahneman (1992) further proposed cumulative prospect theory (CPT) to avoid certain drawbacks inconsistent with first-order stochastic dominance. CPT can account for diminishing sensitivity, loss aversion, and different risk attitudes. Some financial phenomena and the paradoxes of Allais and Ellsberg cannot be explained by traditional theories like EUT. Therefore CPT offering feasible interpretations is widely applied. Benartzi and Thaler (1995) gave explanation for the famous equity premium puzzle using CPT. Because CPT can describe the behavior of bounded rational decision makers in a psychologically more realistic way, over the past decade, researchers in the field of behavioral economics have repeatedly considered how prospect theory should be applied in economic settings. These efforts are now bearing fruit (Barberis, 2013).

CPT has received a great deal of attention, to the best of our knowledge, very few papers have used CPT to solve the problem of portfolio choice. Stracca (2002) considered the optimal allocation of risky assets from identically distributed and symmetric sources under CPT, and Levy and Levy (2004) showed that the mean-variance (MV) and PT-efficient sets almost coincide with each other under the normal distribution assumption, allowing investors to use the MV optimization algorithm to create PT-efficient portfolios. Bernard and Ghossoub (2010) studied the optimal portfolio choice for investors under CPT, and derived some properties of the optimal holding. He and Zhou (2011) developed a new measure of loss aversion, which is the criterion for the well-posedness of the model for large payoffs, and obtained optimal single-period solutions under CPT. Pirvu and Schulze (2012) provided a two-fund separation theorem for risk-free assets and the risky portfolio in which the excess return follows an elliptically symmetric distribution under CPT.

The purpose of this dissertation is to study how CPT decision makers optimize their portfolios. And the effects of certain parameters of CPT about portfolio choices and final returns are analyzed.

1.2 Main objectives of the dissertation

The purpose of this dissertation is to identify potential benefits of behavior based CPT model depending on different market situations in comparison with traditionally accepted portfolio optimisation models. The main objectives are as follows:

1. The appropriate solution approaches were developed to solve the portfolio choice problem under CPT when the joint distribution of portfolio returns is subject to multivariate normal distribution.

2. The appropriate solution approaches were developed to solve the portfolio choice problem under CPT were evaluated when return of each asset has different types of probability distributions.

3. The performances of portfolio choice under CPT with VaR or CVaR constraints and other constraints were investigated in details.

1.3 Thesis Structure

The dissertation comprises of six chapters, a bibliography and appendices.

Chapter 1 is an auxiliary part of the present work that provides background information, my frame work of present study and value together with contributions.

Chapter 2 provides a literature survey for the main theories and mathematical formulations of the considered portfolio optimisation models as well as the definitions of some risk measures, respectively.

Chapter 3 presents an operational model for portfolio selection under CPT and proposes a real-coded genetic algorithm (RCGA) to solve the problem of portfolio choice. To overcome the limitations of RCGA and improve its performance, an adaptive method was developed and a new selection operator was proposed. Computational results show that the new method is a rapid, effective, and stable genetic algorithm.

In Chapter 4 we study the portfolio selection problem under CPT and present a solution for portfolio optimization, the method of coupling scenario generation techniques with a genetic algorithm. Computational results show that the proposed method solves effectively the portfolio selection model. We compare the portfolio choices of CPT investors based on different bootstrap techniques for scenario generation and empirically examine the impact of reference points on investment behavior.

In Chapter 5 we study the portfolio behavior of CPT investor with risk constraints of VaR or CVaR and other constraints.

In Chapter 6 we describe the most important findings and conclusion. The main contribution of this dissertation as well as related future work are presented in this chapter.

Chapter 2

Literature review

2.1 Modern portfolio theory

2.1.1 Mean and Variance analysis

Modern portfolio theory (MPT) provides a mathematical framework assumes that investors act rationally and has preferences in the light of the mean and the variance of returns, which are the random variable. That is investors cannot predict the future return accurately but can predict the expected value of returns according to the probability distribution of returns. According to the probability theory, the expected value is a kind of distributional average.

Except returns, investors also care about the risk. Markowitz (1952) introduced the variance or standard deviation of returns as a risk measurement. The variance of returns reflects squared deviations from the mean so large deviations above or below the mean. The standard deviation of returns is the square root of the variance.

Let R_i denote the return on asset i.

$$R_i = \frac{P_{i,t+1} - P_{i,t}}{P_{i,t}} \tag{2.1}$$

where $P_{i,t}$ and $P_{i,t+1}$ represents the price of asset i at period t and $t+1$ respectively.

Let the return of asset is a bounded discrete random variable and has a probability mass function $g(r_i)$. The expected value of returns is denoted as follows:

$$E(R_i) = \mu_i = \sum r_i g(r_i) \tag{2.2}$$

Let the return of asset is a bounded continuous random variable and has a probability density function $f(r_i)$. The expected value of returns is denoted as follows:

$$E(R_i) = \mu_i = \int_{-\infty}^{+\infty} r_i f(r_i)\, dr_i \tag{2.3}$$

The variance of R_i, denoted by $Var(R_i)$, is defined as follows:

$$Var(R_i) = \sigma_i^2 = E[(R_i - \mu_i)^2] \tag{2.4}$$

When there are n risky assets, the investor wants to apportion his budget to these assets by deciding on a specific allocation $\boldsymbol{x} = (x_1, \cdots, x_n)^T$, $x_i \geq 0$ (i.e. short sales are disallowed) and $\sum_{i=1}^{n} x_i = 1$ (budget constraint) [1]. The multi-assets investment problem is defined as follows:

$$R = \sum_{i=1}^{n} x_i R_i \tag{2.5}$$

The expected value of portfolio return is

$$E(R) = \mu = \sum_{i=1}^{n} x_i E(R_i) = \sum_{i=1}^{n} x_i \mu_i \tag{2.6}$$

The variance of portfolio return is

$$\sigma^2 = Var(R) = \sum_{i=1}^{n} \sum_{j=1}^{n} x_i x_j \sigma_{ij} = \sum_{i=1}^{n} x_i^2 \sigma_i^2 + \sum_{i=1}^{n} \sum_{\substack{j=1 \\ j \neq i}}^{n} x_i x_j \sigma_{ij} \tag{2.7}$$

When $i \neq j$,$\sigma_{i,j}$ represents

$$Cov(R_i, R_j) = E[(R_i - \mu_i)(R_j - \mu_j)] \tag{2.8}$$

When $i = j$,σ_{ii} represents σ_i^2.

Equation (2.7) can be written

$$\sigma^2 = \boldsymbol{x}^T \boldsymbol{\Sigma} \boldsymbol{x} \tag{2.9}$$

where $\boldsymbol{x}$ is a column vector whose components are the x_i, $\boldsymbol{x}^T$ is the row vector that is the transpose of $\boldsymbol{x}$, and $\boldsymbol{\Sigma}$ is the covariance matrix, which is positive definite matrix.

[1] Throughout the dissertation boldface characters denote vectors.

2.1.2 Efficient frontier

The efficient frontier, or portfolio frontier, is the curve that shows all efficient portfolios under frame of MPT introduced by Markowitz (1952). Formally, efficient frontier is the set of portfolios that maximizes the expected return at given risk or minimizes the risk at give expected return.

To obtain the efficient frontier people firstly have to minimize the risk at given expected return μ_0 as follows:

$$\begin{aligned} & \min \sigma^2 = \boldsymbol{x}^T \boldsymbol{\Sigma} \boldsymbol{x} \\ s.t. \quad & \sum_{i=1}^{n} x_i E(R_i) = \mu_0, \\ & \sum_{i=1}^{n} x_i = 1 \\ & x_i \geq 0 \end{aligned} \tag{2.10}$$

Then varying μ_0 between the return on the minimum variance portfolio and the return on the maximum return portfolio traces out the efficient frontier, as shown in the Figure 2.1.

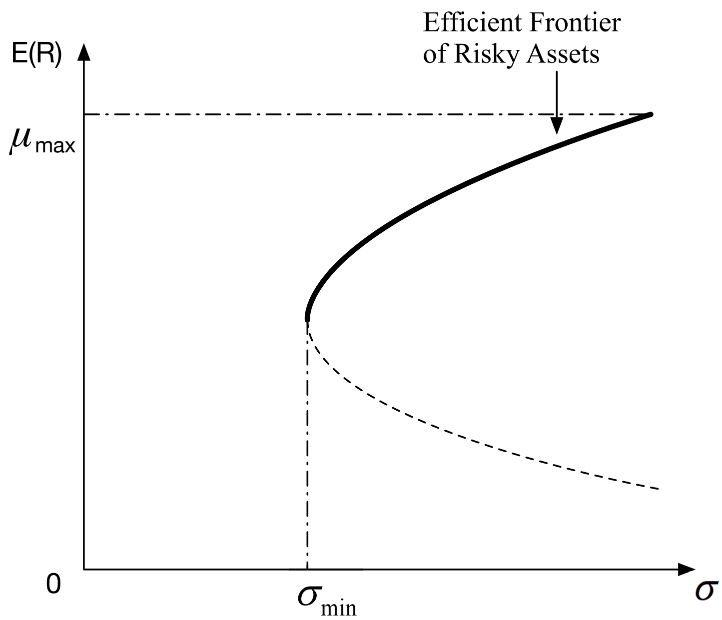

Figure 2.1: The efficient frontier

The efficient frontier can be obtained in a similar way by using the following model:

$$\max E(R) = \sum_{i=1}^{n} x_i E(R_i) \tag{2.11}$$
$$s.t. \qquad \boldsymbol{x}^T \boldsymbol{\Sigma} \boldsymbol{x} = \sigma_0^2,$$
$$\sum_{i=1}^{n} x_i = 1$$
$$x_i \geq 0$$

Investors can choose any point on the efficiency curve according to their attitude toward risk. As mentioned above, an investor can select a point whose has the minimum variance or maximum return on the efficient frontier curve. People also can invest in the portfolio at point S in Figure 2.2 where has maximum return per unit risk.

Take equation (2.11) example, it can be solved by Lagrange multiplier:

$$\boldsymbol{x} = \frac{B(AE_0 - B)}{D} \frac{\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}}{B} + \frac{A(C - BE_0)}{D} \frac{\Sigma^{-1}\mathbf{1}}{A} \tag{2.12}$$

where $\mathbf{1} = (1, ..., 1)^T$, $A = \mathbf{1}^T \Sigma^{-1} \mathbf{1}$, $B = \mathbf{1}^T \Sigma^{-1} \boldsymbol{\mu}$,$C = \boldsymbol{\mu}^T \Sigma^{-1} \boldsymbol{\mu}$,$D = AC - B^2$.

Markowitz (1959) proposed a method to model the willingness to pay (WTP) for risky assets as a tradeoff between their returns and risk, i.e., investors will try to minimize level of risk at given return.

$$WTP(R) = E(R) - bVar(R) \tag{2.13}$$

where b denotes the property of the tradeoff between the maximization of return and minimization of risk and serves as an individual difference index of risk attitude (Glimcher and Fehr, 2013).

The idea of tradeoff between return and risk is widely applied in finance (Sharpe, 1964, Levy and Markowitz, 1979). Some utility functions have interpretation about the tradeoff of between return and risk. And different utility functions have different functional forms for risk (Jia and Dyer, 1996).

The WTP portfolio choice model is, therefore,

$$\begin{aligned} \max_{\boldsymbol{x}} \quad & WTP(R) \\ s.t. \quad & \sum_{i=1}^{n} x_i = 1 \\ & x_i \geq 0 \end{aligned} \tag{2.14}$$

2.1.3 Sharpe ratio

Roy (1952) introduced a method that maximizing the ratio before sharpe ratio was put forward as follow

$$\frac{m-d}{\sigma} \tag{2.15}$$

where m is expected gross return, d represents, in a sense, disaster level and σ is standard deviation of returns.

In finance, Sharpe ratio is a way to test the performance of investment by adjusting for its risk. Sharpe (1966) introduced Sharpe ratio when a risk-free asset to the mix defined as follows

$$Sharpe \quad ratio = \frac{\mu - r_0}{\sigma} \tag{2.16}$$

where r_0 represents the rate of risk-free asset.

Sharpe ratio means the excess return (the expected return over risk-free rate) per unit of risk and is thus the concept of relative value. The portfolio with a higher sharpe ratio provides better return for same risk. Graphically, the maximum Sharpe ratio on the efficient frontier curve is the point M where a line through the r_0 is tangent to the efficient frontier. As shown in Figure 2.3. The line combining r_0 and M is called the capital allocation line (CAL) and represents the combination of the market portfolio and the risk-free asset. The CAL tells investors a truth that they can earn how much excess returns for accepting additional risk.

2.2 Measures of risk

2.2.1 Variance and Semivariance

The concept of risk is an important factor to be considered in the portfolio selection problem. Markowitz proposed to measure the risk via the deviation from the mean, i.e.,

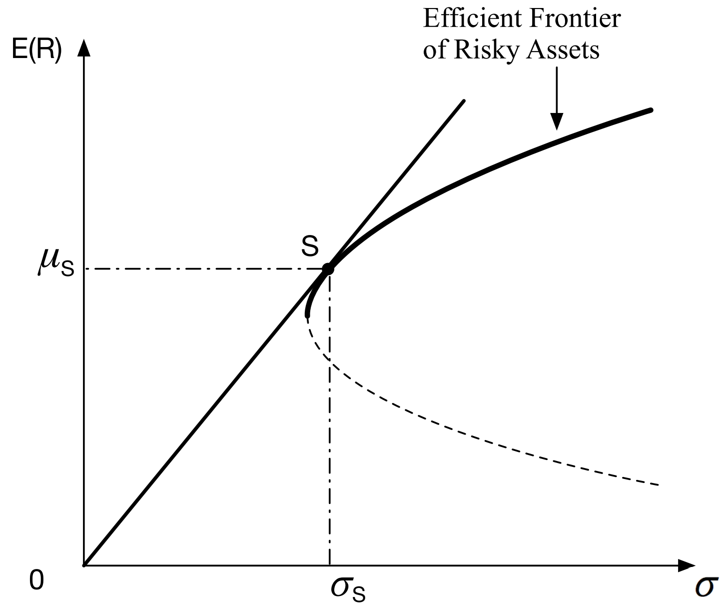

Figure 2.2: The Sharpe ratio

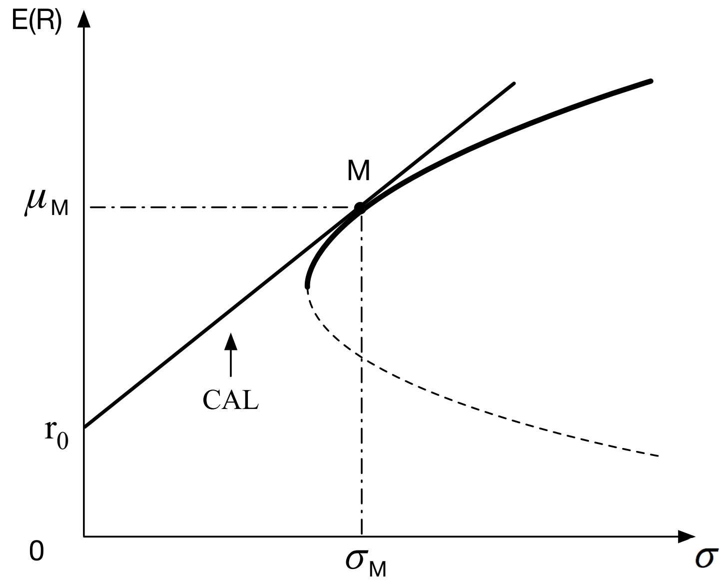

Figure 2.3: Capital market line

variance or standard deviation. By using covariance between all pairs of risky assets, risk level can be measured. The major contribution by Markowitz is to measure the risk of a portfolio by means of joint distribution of returns of all assets and this is the first mathematical formalization of the idea of diversification of investments for reducing risk [2]. The variance, as a measure of risk, has an advantage of simplicity that it is very important to portfolio selection problem. However, there are some criticism about variance as a risk measurement.

Variance measures the upside and downside risk as well. Downside risk is the risk being below the expected return, which is the financial risk associated with losses. Generally, people do not treat gains and losses equivalently and pay more attention to downside risk (Horcher, 2011, Nawrocki, 1999). Therefore some scholars suggested that semivariance should be a tool to measure risk (Porter, 1974, Hogan and Warren, 1974, Estrada, 2007, Jin et al., 2006). Semivariance here refers to downside risk rather than upside risk. Mathematically, the semivariance is expressed as follow:

$$SVar(R) = E[(R - E(R))^2 \mathbb{1}_{\{R \leq E(R)\}}] \tag{2.17}$$

where $\mathbb{1}_{\{R \leq E(R)\}}$ is an indicator function.

Variance is the concept of second moment and may ignore the risk from the higher moments of the probability distribution. The model of Markowitz is applied only to the case of elliptic distributions, such as normal or t-distributions with finite variances, which seems somewhat different from reality. It has been proved that the distributions of returns on many risky assets are skewed, leptokurtic, and heavy tail [3].

2.2.2 VaR

VaR has become one of the standard instruments to measure risk for both banks and other financial institutions. Regulators such as the Bank for International Settlements recommend VaR-measures to determine capital adequacy requirements. Generally, investors seldom allow their potential loss to exceed a certain level. Consequently, VaR is used as a risk measurement on portfolio problem to control the risk. VaR has become an industry standard for risk measurement and percentile based indicator (JPMorgan, 1994). It is usually defined as the worst loss over a target horizon that will not be

[2] It is noteworthy that Irving (1906) has firstly suggested to use variance as a measure of economic risk. Marschak (1938) suggested using the means and the covariance matrix of consumption of commodities as a first order approximation in measuring utility.

[3] Multivariate normal distribution is attractive because the association between any two random variables can be expressed by their marginal distribution and correlation coefficient.

exceeded with a given level of confidence (Jorion, 2006).

For a target horizon and given a confidence level $c \in (0,1)$, the VaR at confidence level c is the smallest value l such that the probability that the loss L exceeds l is no larger than $(1-c)$. Mathematically, VaR can be expressed as follow:

$$VaR_c(\boldsymbol{x}, \boldsymbol{R}) = inf\{l \in \mathbb{R} : P(L > l) \leq 1 - c\} \tag{2.18}$$

L is a random loss variable with the cumulative distribution function as follow:

$$F_L(l) = P(l \leq l) \tag{2.19}$$

Some scholars consider that VaR is a natural progression from MPT in some aspects(Dowd, 2002). However, there are some differences as follows:

- MPT explains the risk according to the mean and standard deviation of returns, whereas VaR interprets risk in terms of the maximum likely loss.
- MPT assumes that the returns follow normal distribution, while VaR can work at a wider range of possible distributions.
- MPT responds to only market risks, while VaR can be used not only for market risk, but also for credit, liquidity and other risks.

There is growing interest in VaR and for various applications including financial institutions, regulators, nonfinancial corporations, and asset managers. Any institution, which is susceptible to risk, can use the VaR to report risk information and to control risk even manage risk. Institutional investor, such as Chrysler pension fund are now turning to VaR to manage their financial risk. The Basel Committee on Banking Supervision,the U.S. Federal Reserve, the U.S. Securities and Exchange Commission, and regulators in the European Union has converged on VaR as benchmark risk measure(Jorion, 2006).

According to equation (2.18)，an investor can control the risk for minimizing the VaR with constraint of expected return as follows:

$$\begin{aligned} \min_{\boldsymbol{x}} \quad & VaR_c(\boldsymbol{x}, \boldsymbol{R}) \\ s.t. \quad & \boldsymbol{x}^T E(\boldsymbol{R}) \geq \mu_0, \\ & \boldsymbol{x}^T \mathbf{1} = 1 \end{aligned} \tag{2.20}$$

where $\mathbf{1} = (1, 1 \ldots, 1)^T$, μ_0 represents the requested expected return.

Some scholars studied the problem of maximizing the expected utility under a VaR constraint and mean-VaR model has been proposed (Campbell et al., 2001, Alexander and Baptista, 2002, Consigli, 2002). Until now, the mean-VaR model, a method to decide optimal selections in terms expected return and VaR, remains as an important research subject of portfolio selection under riskTsao (2010), Sheng et al. (2012), Ali and Jilani (2014). The mean-VaR model can be given as follow:

$$\begin{aligned} \max_{\boldsymbol{x}} \quad & \boldsymbol{x}^T E(\boldsymbol{R}) \\ s.t. \quad & VaR_c(\boldsymbol{x}, \boldsymbol{R}) < l \\ & \boldsymbol{x}^T \mathbf{1} = 1 \end{aligned} \tag{2.21}$$

where $\mathbf{1} = (1, 1 \ldots, 1)^T$.

2.2.3 CVaR

VaR is a popular risk measurement although some of its mathematical properties have influence on application for optimal portfolio problem, and can not respond to the magnitude of the possible losses below the threshold it identifies.

Generally, the criticism of VaR is manifested mainly in three aspects. Firstly, VaR only measures the most loss if the tail event does not occur, i.e., VaR fails to provide information beyond the tail of distribution which may be exposed to the danger of a very large loss. Secondly, VaR sometimes contradicts the sub-additivity property of coherent risk measure, which is proposed by Artzner et al. (1999). This means that aggregating individual risks do not increase the overall risk[4]. Thirdly, it is difficult to optimize the portfolio problem based on VaR if the returns or losses are specified according to the scenarios. In fact, VaR function is non-smooth and non-convex with respect to the portfolio ratio $\boldsymbol{x}$ and exhibits multiple local extrema (Topaloglou et al., 2002).

Some scholars suggested that conditional value-at-risk (CVaR) is an alternative percentile measure of risk(Pflug, 2000, Rockafellar and Uryasev, 2000, 2002). CVaR, unlike VaR, can quantify the losses beyond VaR. CVaR is defined as the conditional expectation of losses exceeding VaR at a given confidence level. CVaR at confidence level $c \in (0, 1)$ for loss L of a portfolio is defined to be

$$CVaR_c = E(L | L \geq VaR_c) \tag{2.22}$$

[4]If the distribution is elliptical, then VaR is a coherent measure.

According to equation (2.22), an investor can control the risk for minimizing the CVaR with constraint of expected return as follows:

$$\begin{aligned} \min_{\boldsymbol{x}} \quad & CVaR_c(\boldsymbol{x}, \boldsymbol{R}) \\ s.t. \quad & \boldsymbol{x}^T E(\boldsymbol{R}) \geq \mu_0, \\ & \boldsymbol{x}^T \mathbf{1} = 1 \end{aligned} \tag{2.23}$$

where $\mathbf{1} = (1, 1 \ldots, 1)^T$, μ_0 represents the requested expected return.

Similarly, the mean-$CVaR$ model can be given as follow:

$$\begin{aligned} \max_{\boldsymbol{x}} \quad & \boldsymbol{x}^T E(\boldsymbol{R}) \\ s.t. \quad & CVaR_c(\boldsymbol{x}, \boldsymbol{R}) < l \\ & \boldsymbol{x}^T \mathbf{1} = 1 \end{aligned} \tag{2.24}$$

where $\mathbf{1} = (1, 1 \ldots, 1)^T$.

Convexity can be preserved in the case of optimizing problems. The random variables are discrete under various scenarios, then CVaR optimisation can be expressed as a linear programming(Rockafellar and Uryasev, 2002).

2.3 Expected utility theory

2.3.1 Expected utility theory

In the mid-seventeenth Century, people began to use the expected value theory to consider gambling problem, which was the problem of maximization of expected value of gamble X as follows:

$$EV(X) = \sum_i p_i x_i \tag{2.25}$$

However, St. Petersburg paradox has suggested that human behaviors may significantly deviate from expected value theory. According to the paradox, people were unwilling to pay too much money for a gamble in which they can get 2^n ducats when the coin lands "head" on the ground for the first time at the n-th throw. Thus it is noteworthy that the gamble has infinite expected value as follow:

$$E = \frac{1}{2} \cdot 2 + \frac{1}{4} \cdot 4 + \cdots + \frac{1}{2^n} 2^n + \cdots = 1 + 1 + \cdots = \infty \tag{2.26}$$

The original idea of expected utility theory was first proposed by Daniel Bernoulli in 1738 to solve St. Petersburg paradox by using expected utility instead of expected value (Bernoulli, 1954). As a decision model under risk, expected utility theory has attracted wide attention until Neumann et al. (1944) suggested that the theory could be explained systematically by a set of axioms on preferences.

Suppose that people are faced with a choice between two outcomes, A_1 and A_2. $A_1 \succ A_2$ means that A_1 is strictly preferred to A_2, that is, people are willing to select the A_1 when A_1 and A_2 are offered. $A_1 \sim A_2$ means that people evaluate the two outcoms the same, i.e. indifference. $A_1 \succeq A_2$ means that people prefer A_1 or are indifferent between A_1 and A_2.

Neumann et al. (1944) proposed the four axioms as follows:

Axiom 1 (Completeness): For all A_1, A_2, exactly one of the following holds: $A_1 \succ A_2$, $A_2 \succ A_1$, or $A_1 \sim A_2$.

Axiom 2 (Transitivity): If $A_1 \succeq A_2$, $A_2 \succeq A_3$, then $A_1 \succeq A_3$.

Axiom 3 (Continuity): $A_1 \succeq A_2 \succeq A_3$, and there exists a probability $p \in [0, 1]$, then $pA_1 + (1 - p)A_3 \sim A_2$.

Axiom 4 (Independence): If $A_1 \sim A_2$, for any A_3 and $p \in [0, 1]$, $(A_1, A_3, p) \sim (A_2, A_3, p)$

Axiom 5 (von Neumann–Morgenstern utility theorem): If Axiom 1-4 are satisfied, there exists a function $u(\cdot)$ assigning to each outcome A_i a real number such that,

$$\begin{aligned} A_1 \succeq A_2 \quad iff \quad u(A_1) \geq u(A_2), \\ u(A_1, A_2, p) = pu(A_1) + (1 - p)u(A_2) \end{aligned} \tag{2.27}$$

Pennacchi (2007) has given further interpretation of axioms above and proof of how they lead to the von Neumann and Morgenstern EU decision rule.Ingersoll (1987), Huang and Litzenberger (1988) and Levy (2011) also provided some similar descriptions. Here, we refer to von Neumann–Morgenstern utility function simply as the expected utility function (EUT).

EUT was developed by John von Neumann and Oskar Morgenstern in an attempt to define rational behavior when people face uncertainty. This theory contends that individuals should act in a particular way when confronted with decision-making under uncertainty. In this sense, the theory is "normative," which means that it describes how people should rationally behave. This is in contrast to a "positive" theory, which characterizes how people actually behave.

Indeed, EUT has given the simplicity and numerical modeling tool. Consequently, their axioms have made utility theory a powerful tool for studying the decision making behavior.

2.3.2 Risk attitudes

It is believed that people have different attitudes to risk: risk-averse, risk-seeking and risk-neutral. The utility function is useful in defining risk preferences. Consider the prospect

$$(a_1, p; a_2, q) \tag{2.28}$$

where outcomes $a_1, a_2 \geq 0$ and probability $q = 1 - p$.

People's preferences can be described by relationship between the utility of the expected value of a prospect and the expected utility of the prospect, as shown in the Figure 2.4 .

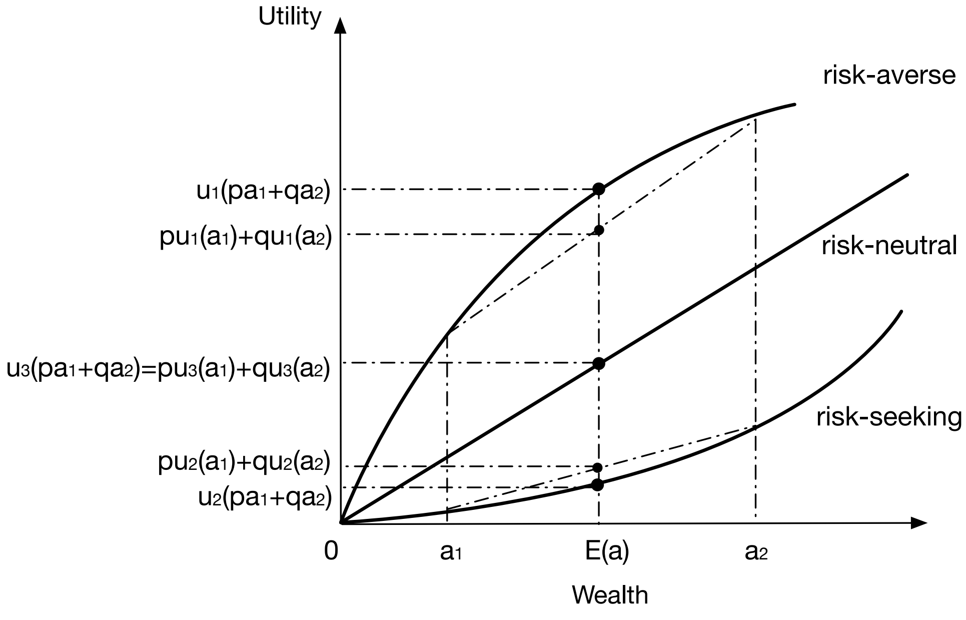

Figure 2.4: Utility functions and risk attitude.

If

$$u_1(pa_1 + qa_2) > pu_1(a_1) + qu_1(a_2) \tag{2.29}$$

then the person is type of risk-averse.

If

$$u_2(pa_1 + qa_2) < pu_2(a_1) + qu_2(a_2) \tag{2.30}$$

then the person is type of risk-seeking.

If

$$u_3(pa_1 + qa_2) = pu_3(a_1) + qu_3(a_2) \tag{2.31}$$

then the person is type of risk-neutral.

For EUT, decision makers' attitudes towards uncertainty are wholly modeled by the value of utility functions defined on final asset positions. Every rational decision maker is assumed to make decisions following the principle of maximizing the value of his expected utility. The expected utility of a choice is the sum of the utility functions of possible N outcomes weighted by the corresponding probabilities:

$$\sum_{i=1}^{N} p_i u(x_i) \tag{2.32}$$

Von Neumann and Morgenstern stated in their expected utility theory that the utility function exists if and only if the preferences of an individual satisfy **Axioms 1-Axioms 4**.

For EUT, the utility function is assumed to be concave, which means the diminishing marginal utility obtained from an extra unit of return. The degree of risk aversion is captured by the shape of the utility function. For decades, the EUT played a dominant role in the decision making problems in various areas of economics.

2.3.3 Portfolio Selection Problem under EUT

EUT has been used as a reference to find the optimal solution in many areas of economics, as a result, a decision is a choice between some subset of all possible states and all feasible weighted portfolios. Each investment choice reduces to a prospect with a probability distribution. The rule of decision making is to maximize $E[u(\cdot)]$, where $u(\cdot)$ is a real valued function representing the utility obtained from certain wealth or returns. Generally, the usual assumption in EUT is that decision makers are risk-averse, which means that the $u(\cdot)$ is an increasing concave (i.e., $u'(\cdot) > 0$ and $u''(\cdot) < 0$).

Suppose that investors are faced with one of two possible investments, each of them has n consequences, denoted by $S_1, S_2, \ldots, S_n$. Suppose that the first investment will result in n consequences with $p_i, i = 1, 2, \ldots, n$ whereas the second one will produce probability $q_i, i = 1, 2, \ldots, n$. The object of decision maker can be described as follows:

$$\max \quad \{E_1(u(S_i)), E_2(u(S_i))\} \tag{2.33}$$

where $E_1(u(S_i))$ and $E_2(u(S_i))$ represent $\sum_{i=1}^{n} p_i u(S_i)$ and $\sum_{i=1}^{n} q_i u(S_i)$ respectively.

Different investment proportions will result in different probability distribution of portfolio returns. People invest x_i dollars in the i-th risky asset and $(W_0 - \sum_{i=1}^{n} x_i)$ dollars in the risk-free asset, the final wealth in the next period can be denoted by $\overline{W}$ as follows:

$$\overline{W} = W_0(1 + r_0) + \sum_{i=1}^{n} x_i(r_i - r_0) \tag{2.34}$$

where W_0 is initial wealth, r_0 represents rate of risk-free asset, r_i is the random rate of return on the i-th risky asset.

The portfolio choice problem that maximizes the expected utility of one's final wealth in the next period can be expressed mathematically as follows:

$$\max_{x_i} \; E[u(\overline{W})] \tag{2.35}$$

The expected utility function of investor who has an utility $u(R)$ can be defined as:

$$E[u(\overline{W})] = \int_{-\infty}^{+\infty} u(w) dF_{\overline{W}}(w) \tag{2.36}$$

where $F_{\overline{W}}$ is the probability distribution function of $\overline{W}$.

According to equation (2.36), if different investment decisions have the same distribution function, they will produce the same expected utility and are indifferent to each other.

2.4 Prospect theory

Although MPT, VaR or CVaR, and EUT have been very useful in modeling portfolio choice, substantial empirical and experimental evidence has revealed that they do not reflect reality. To solve the problem, the original version of prospect theory was proposed by Kahneman and Tversky (1979).

PT (prospect theory) is able to find a solution to several paradoxes in decision theory under uncertainty like reported by Allais and Ellsberg in which people's choices violate the postulates of subjective expected utility.

The PT models includes two-stages: the first stage involves editing, and the second involves evaluation. The use of an editing phase is the most obvious distinguishing characteristic of PT from any of the theories discussed in the previous section. The editing phase is the most obvious feature to distinguish from any other theory mentioned above.

There are four important differences between EUT and PT in terms of decision making: reference dependence, different risk attitude, loss aversion, and probability distortion. These are the factors which make PT psychologically more realistic.

First, PT investors appraise their investment according to its relative value with respect to some reference point, which separates the investment into gains and losses. In contrast, EUT implies that investors make choices based on changes for final value.

Second, PT investors display different behaviors with respect to gains and losses. As such, as shown in Fig.2.5, the value function is concave with respect to gains and convex with respect to losses. The concavity over gains reflects the finding that people tend to be risk averse over moderate probability gains: they typically prefer a certain gain of $1000 to a 50 percent chance of $2000. However, people also tend to be risk seeking over losses: they prefer a 50 percent chance of losing $2000 to a guarantee of losing $1000. EUT is typically concave everywhere, i.e. risk averse.

Third, PT investors are more sensitive to losses than to gains of the same magnitude, i.e. loss averse. Loss aversion, as an important concept in prospect theory, implies that the utility function is steeper for losses than it is for gains, as shown in Fig. 2.5. Loss aversion indicates that most people are unwilling to take part in a gamble consisting of a 50 percent chance of losing $1000 and a 50 percent chance of gaining $1100. Benartzi and Thaler (1995) discussed an equity puzzle and concluded that, if loss aversion is taken into account, the risk premium can be more substantial than when it is not considered. Thaler (1980) discussed the endowment effect using loss aversion, concluding that people value their own stuffs more than those to others. Samuelson and Zeckhauser (1988) discussed the status quo bias whereby most real decision-makers prefer to maintain their current or previous decisions because of loss aversion. Barberis et al. (2006) discussed the stock market nonparticipation phenomenon in which, even though the stock market has a high mean return and a low correlation with other household risks, many households have historically been reluctant to allocate any money to it because of loss aversion. There is no concept of loss aversion in EUT, therefore the explanation of the above phenomena is beyond its scope.

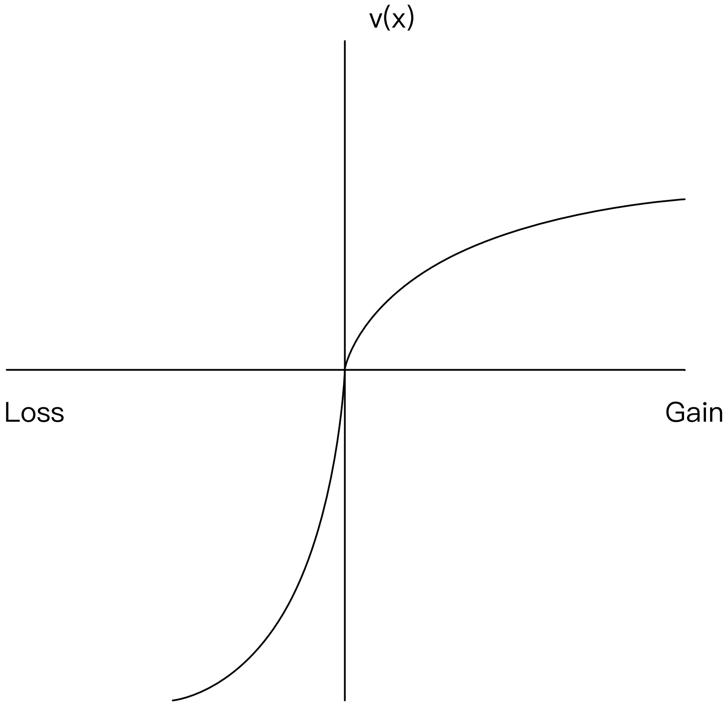

Figure 2.5: The value functions.

Finally, PT investors do not weight outcomes using objective probabilities, unlike EUT, but rather by transformed probabilities obtained via a probability weighting function, as shown in Figure 2.6. The paradoxes of Allais and Ellsberg can be explained by means of a nonlinear transformation of the objective probabilities.

Consider the prospect[5]

$$(t_1, p; t_2, q) \tag{2.37}$$

to be read as gain t_1 with probability p and t_2 with probability q, where $t_1, t_2 \neq 0$ and $p + q = 1$. In the original version of prospect theory, the agent assigns the prospect the value

$$w(p)v(t_1) + w(q)v(t_2) \tag{2.38}$$

where $v(\cdot)$ and $w(\cdot)$ are known as the value function and the probability weighting function, respectively. These functions satisfy $v(0) = 0$, $w(0) = 0$, and $w(1) = 1$.

[5]It is noteworthy that PT can be applied only to gambles with at most two nonzero outcomes (Kahneman and Tversky, 1979, Barberis, 2013).It is wrong to use PT to solve more than two nonzero outcomes by some scholars.

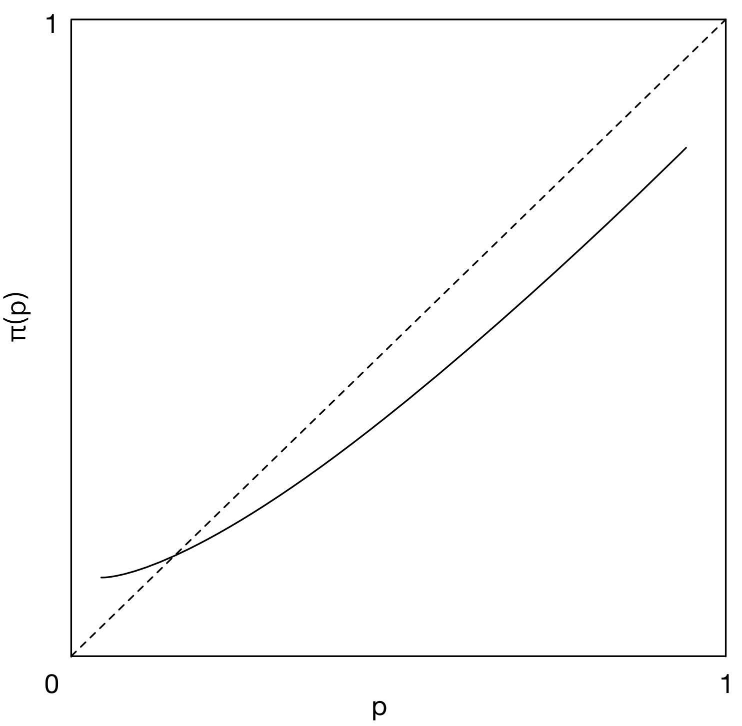

Figure 2.6: The weighting functions.

2.5 Cumulative Prospect theory

Cumulative prospect theory (CPT) is a modified version of prospect theory proposed by Tversky and Kahneman (1992), who put forward explicit functional forms for $v(\cdot)$ and $w(\cdot)$ and applied the probability weighting function to the cumulative probability, not to the single probability. This ensures that CPT does not violate first-order stochastic dominance—a weakness of the original prospect theory—and that it can be applied to gambles with any number of outcomes, not just two.

Moreover, Tversky and Kahneman drawed a conclusion to the important "four-fold pattern of risk attitudes", which is risk-seeking for small-probability gains and large-probability losses and risk-aversion for small-probability losses and large-probability gains. This can explain why people like both lotteries and insurance, which are difficult to rationalize under EUT.

According to Tversky and Kahneman (1992), the CPT investors evaluate the investment[6]

$$(t_{-m}, p_{-m}; \ldots; t_{-1}, p_{-1}; t_0, p_0; t_1, p_1; \ldots; t_n, p_n) \tag{2.39}$$

[6] $t_{-m}, ..., t_n$ are the results that the actual outcome minus the value of reference point.

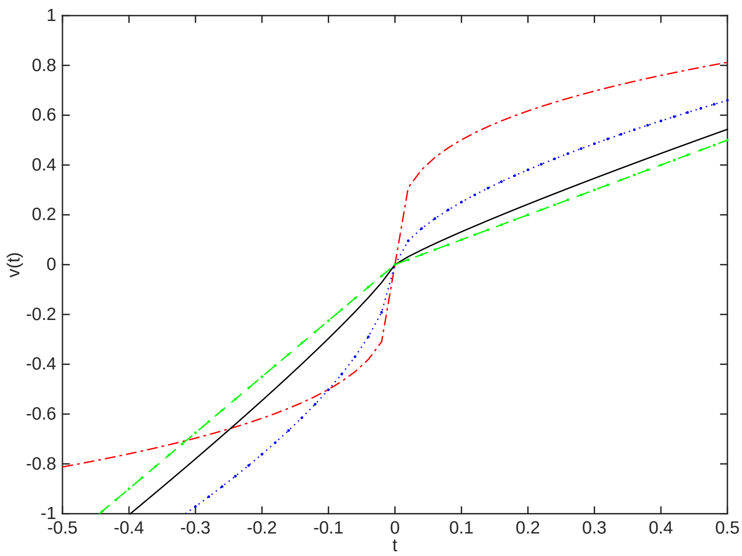

Figure 2.7: The value functions

where $t_i < t_j$ for $i < j$, $t_0 = 0$ and $\sum_{i=-m}^{n} p_i = 1$.

As mentioned above, EUT assumes that the investors are risk-averse in the gains. However, CPT assumes that investors express the outcomes as deviations from some reference point and response being more sensitive to losses than to gains. The value function $v(\cdot)$ is defined by Tversky and Kahneman (1992) as:

$$v(t) = \begin{cases} t^{\alpha} & t \geq 0 \\ -\lambda(-x)^{\beta} & t < 0 \end{cases} \tag{2.40}$$

where $\alpha = \beta = 0.88$ and $\lambda = 2.25$. [7]

For $\alpha, \beta < 1$, the S-shaped power value function exhibits risk aversion over gains and risk seeking over losses. The parameter λ captures loss aversion, assuming that investors consider losses to be more than twice as important as gains. The value functions $v^{+}(\cdot)$ and $v^{-}(\cdot)$ are often supposed to be an increasing, twice differentiable, invertible, and concave functions (Bernard and Ghossoub, 2010).

The parameters of value function define the degree of risk aversion with respect to gains, the degree of risk seeking with respect to losses, and the degree of loss aversion. The parameter α represents risk aversion with respect to gains and the parameter β represents risk preference with respect to losses. The parameter λ represents the loss aversion: the higher the value of λ, the more loss-averse the CPT investors. As shown in Figure 2.7, the dash-dot curve corresponds to $\alpha = \beta = 0.3$, $\lambda = 1$; the dotted curve

[7]The value functions in PT and in CPT are often confused. Kahneman and Tversky (1979) only described the form of value function whereas the power value functions and their parameters were proposed by Tversky and Kahneman (1992).

corresponds to $\alpha = \beta = 0.6$, $\lambda = 2$; the solid curve corresponds to $\alpha = \beta = 0.88$, $\lambda = 2.25$; the dashed line corresponds to $\alpha = 1$, $\beta = 1$, $\lambda = 2.25$.

CPT investors do not weight the outcomes according to objective probabilities. Moreover, the weighting functions have different parameters over the domains of gains and losses, denoted by $w^+(\cdot)$ and $w^-(\cdot)$, respectively. Tversky and Kahneman (1992) proposed the following functions:

$$w^+(P) = \frac{P^\gamma}{[P^\gamma + (1-P)^\gamma]^{1/\gamma}} \tag{2.41}$$

$$w^-(P) = \frac{P^\delta}{[P^\delta + (1-P)^\delta]^{1/\delta}} \tag{2.42}$$

Tversky and Kahneman (1992) estimated that $\gamma = 0.61$, $\delta = 0.69$ [8].

The parameters of weighting functions define the degree of distortion to the objective probabilities. The smaller the values of γ, δ, the greater the degree of distortion. As shown in Figure 2.8, the dotted curve corresponds to $\gamma = 0.61$; the solid curve corresponds to $\gamma = 0.69$; the solid line corresponds to $\gamma = 1$.

Some scholars considered that $w^+ : [0,1] \to [0,1]$ and $w^- : [0,1] \to [0,1]$ are non-decreasing and differentiable with $w^+(0) = w^-(0) = 0$ and $w^+(1) = w^-(1) = 1$ (Bernard and Ghossoub, 2010). Ingersoll (2008) showed that $0.28 < \gamma, \delta < 1$ ensures that $w^+(\cdot)$ and $w^-(\cdot)$ are all increasing functions. For the case of $\gamma = \delta = 1$, the weighting functions have the following linear form:

$$w^+(P) = w^-(P) = P \tag{2.43}$$

The decision weights were defined by Tversky and Kahneman (1992) under equation (2.39) as follows:

$$\pi_i = \begin{cases} \pi_i^+ = w^+(p_i + \cdots + p_n) - w^+(p_{i+1} + \cdots + p_n) & 0 \le i \le n \\ \pi_i^- = w^-(p_{-m} + \cdots + p_i) - w^-(p_{-m} + \cdots + p_{i-1}) & -m \le i < 0 \end{cases} \tag{2.44}$$

where $\pi_i^+(\cdot)$ and $\pi_i^-(\cdot)$ are the weighting functions for gains and losses, respectively.

[8] It is incorrect that some papers about the application of PT used equation (2.41) and equation (2.42). No equations and parameters about weighting functions were given by Kahneman and Tversky (1979).

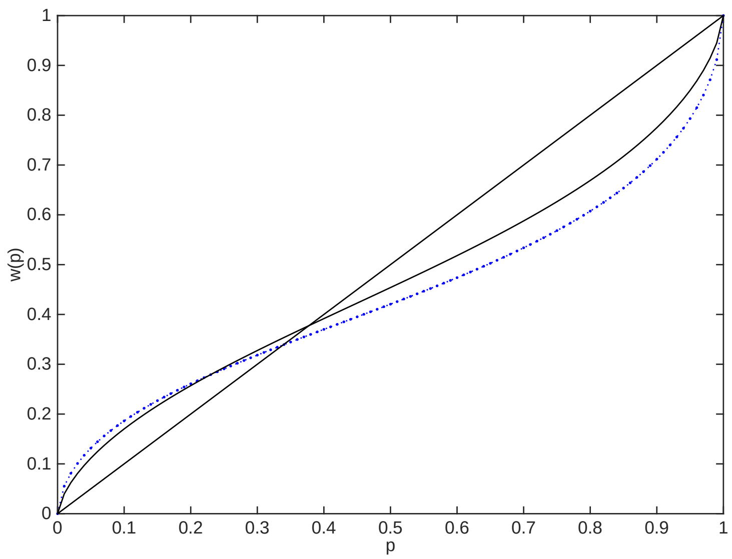

Figure 2.8: The probability weighting functions

The CPT value of the investment for stocks is given by

$$V = \sum_{i=-m}^{n} \pi_i \cdot v(t_i) \tag{2.45}$$

CPT investors make portfolio choices by maximizing their CPT value; that is, CPT investors determine their investments by maximizing the value of equation (2.45).

Tversky and Kahneman (1992) provided an illustration of CPT model but the result was not given. According to equation (2.40), (2.41), (2.42), (2.44) and (2.45), this dissertation gives the result and provides a brief analysis. Assume that there is a zero level of CPT, i.e. CPT value is 0, which is a critical value of taking part in or not taking part in gamble. With the parameters of $\alpha = \beta = 0.88, \gamma = 0.61, \delta = 0.69$ and $\lambda = 2.55$, the CPT value is -1.85. Furthermore, the CPT value is 0.27 when the parameters are $\alpha = \beta = 0.88, \gamma = 0.61, \delta = 0.69$ and $\lambda = 1$. People who have the parameters are proposed by Tversky and Kahneman (1992) will not play the game, whereas, other things being equal, people who do not have the loss aversion, i.e. $\lambda = 1$, will participate in the game.

2.6 Summary

In general, the portfolio problem is how many assets in the portfolio will be necessary, and the optimal level of diversification should provide convenience for the portfolio management.

The application of the MPT makes it possible to solve portfolio problem under return-risk space, i.e., investors face a trade-off between risk and expected return. The efficient frontier represents portfolios which maximize expected return for a given level of risk. No rational investor would therefore choose a portfolio below the curve because a portfolio with higher return and the same risk exists. As capital allocation line (CAL) gives combinations of the risk-free asset and the market portfolio, rational investors should choose a portfolio on this line in order to achieve the highest return with the lowest risk. Widely used measures of risk are value at risk (VaR) and conditional value at risk (CVaR). The former measures the loss value to be exceeded with a specified probability. The latter was used to quantify the average of the losses that occur beyond the VaR cutoff point at a given confidence level in the distribution.

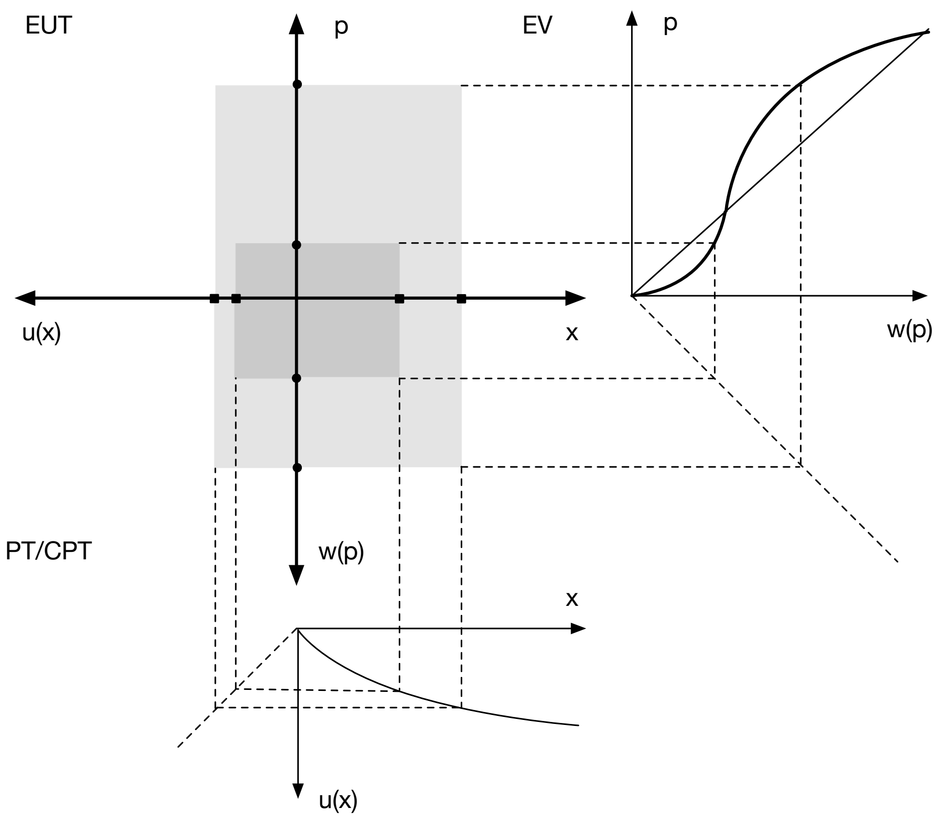

Figure 2.9: EV, EUT and PT/CPT

EUT describes the subjective sensitivity to the results, but treats probability of these results objectively. Lopes (1987) has said that risk attitude is more than the psychophysics of money. Psychologists and economists have different attitudes when they face a same prospect. Psychologists Economists seem to consider that risk is a function of money or return. Psychologists have showed that risk is a feel about probability instead of money or return. It is worth mentioning that the depth of the rank-dependent idea, a transform of cumulative probability instead of individual probability,has taken analysis about risk to the next level.

Finally, this dissertation suggests that EV, EUT, PT and CPT drops in a continuous line to some extent, i.e., the pattern of inner product. There exist two mappings for outcomes and their probabilities, i.e., $u(\cdot)$ and $w(\cdot)$ [9]. As shown in Figure 2.9, the value of EV is the area sum of dark shadow and light shadow at First Quadrant [10]. The Weighted Monetary Value, the products of outcomes and transformed probabilities, is merely the area sum of dark shadow and light shadow at Fourth Quadrant [11]. The EUT, which is the products of transformed outcomes and probabilities, is the area sum of dark shadow and light shadow at Second Quadrant. PT/CPT, which is the products of transformed outcomes and transformed probabilities, is the area sum of dark shadow and light shadow at Third Quadrant[12].

[9] Although utility function in EUT differs from value function in PT/CPT, without loss of generality, they can be assumed to be same thing here.

[10] In fact, either dark shadow or light shadow in each quadrant is rectangle. Part of light shadow is hidden by dark shadow.

[11] Schoemaker (1982) has given various forms of models under the pattern of inner product in detail.

[12] When there are only two non-zero outcomes in prospect, the PT is consistent with CPT. And the value of reference point is assumed to be zero in Figure 2.9.

Chapter 3

Portfolio choice under multivariate normal distribution

In this chapter, we consider the problem of identifying the optimal portfolio consisting of one riskless asset and multiple risky assets under CPT. The CPT function is generally non-convex, non-concave and non-smooth, which means that traditional optimization methods such as Lagrange multipliers and convex duality do not work and the CPT function may have many local maxima (He and Zhou, 2011).

The problem mentioned above can be solved by some stochastic search algorithms, such as genetic algorithm, simulated annealing algorithm, tabu search algorithm and particle swarm algorithm. Stochastic search algorithms offer a number of advantages over more traditional optimization methods. Genetic algorithm and simulated annealing are powerful approaches and well accepted among these methods (Spall, 2005).

Genetic algorithm developed by Holland (1975), is an effective solution to solve problems of optimization. Simulated annealing algorithm, was proposed by Kirkpatrick et al. (1983), is also an effective method for optimization problem. Genetic algorithm has ability of global searching; whereas it is poor in local searching. This algorithm has a problem of premature convergence to be trapped in local optima. Simulated annealing algorithm can find the local optimum solutions quickly and avoid being entrapped into local optimum solutions. However, it is incapable in global searching, leading to infinite time for finding a global optimum. Considering these reasons, we use genetic algorithm to find optimum solution in this dissertation.

Genetic algorithms (GAs) are of robust search and optimization techniques. Unlike gradient-based methods, GAs do not use any properties of the function being optimized. The only requirement of the problem is that objective functions can be computed. Moreover, GAs are superior to gradient-based methods as the search is not biased towards the locally optimal solution. In addition, GAs differ from random sampling algorithms in their ability to direct the search towards relatively “prospective” regions in the search space.

GAs have been successfully used in a variety of research fields. In recent years, numerous studies have showed that GAs can efficiently solve optimal portfolio problems in finance. For instance, Chang et al. (2000) and Yang (2006) used GAs to solve the problem of mean-variance portfolio optimization, and Tsao (2010), Baixauli-Soler et al. (2011), and Ranković et al. (2014) solved optimality problems related to the mean-VAR using GAs. However, to the best of our knowledge, no studies have yet used CPT with GAs to solve the portfolio choice problem save for Grishina et al. (2017).

The main contributions of this chapter are to describe a model that identifies an investment portfolio with several risky assets and one riskless asset under CPT, and to adapt a genetic algorithm to solve the portfolio selection model. We employ a real-coded genetic algorithm (RCGA), which is more consistent, precise, and converges faster than ordinary GAs. The RCGA is therefore better suited to large-dimensional search spaces than a binary-coded genetic algorithm (BCGA) (Baskar et al., 2001, 2003, 2004). However, recent studies have shown that several drawbacks of RCGAs reduce their search capabilities. To enhance the efficiency of our RCGA research, we incorporate some adaptive properties. This dissertation suggests that the resulting adaptive RCGA (ARCGA) is a highly efficient and effective algorithm for obtaining near-optimal solutions within a few minutes.

3.1 Objective functions of CPT investors

In CPT, the probability distribution is assumed to be discrete by Tversky and Kahneman (1992), although we can extend the scope of CPT to more general distributions by adopting the relevant methods (Barberis and Huang, 2008, Jin and Yu Zhou, 2008, Bernard and Ghossoub, 2010).

Assumption 1: CPT investors are more concerned with return than the final wealth.

Let R_p be the rate of return of a portfolio at the end of the period and r_f be the reference point, which separates gains and losses. Define the deviation D from the reference level by

$$D = R_p - r_f \tag{3.1}$$

The cumulative distribution function of the random variable D is denoted by F_D, and $|E(D)| < \infty$ and $Var(D) < \infty$.

As mentioned earlier, CPT treats outcomes as gains and losses from a reference level separately. The value function v is defined as follows:

$$v(t) = \begin{cases} v^+(t) & t \geq 0 \\ -v^-(-t) & t < 0 \end{cases} \tag{3.2}$$

where $v^+ : \overline{\mathbb{R}}^+ \to \overline{\mathbb{R}}^+$ and $v^- : \overline{\mathbb{R}}^- \to \overline{\mathbb{R}}^-$ are invertible, twice differentiable, increasing functions with $v(0) = v^+(0) = v^-(0) = 0$, $v^+(+\infty) = +\infty$ and $v^-(-\infty) = -\infty$.[1] t are the possible values of D.

The value function (3.2) exhibits a S-shaped curve that captures the features described above: it is concave for gains and convex for losses, being more sensitive to losses than to gains, which is so called "loss aversion" in CPT. There are several types of value functions, such as the power value function, exponential value function, quadratic value function, and linear value function. We consider the following two types:

(1) The piecewise power value function

This is the original function considered by Tversky and Kahneman (1992) as the general form of the value function in CPT.

$$v(t) = \begin{cases} t^\alpha & t \geq 0 \\ -\lambda(-t)^\beta & t < 0 \end{cases} \tag{3.3}$$

where $0 < \alpha, \beta \leq 1$ and $\lambda \geq 1$. We call this the general value function.

The parameter α represents risk aversion with respect to gains and the parameter β represents risk preference with respect to losses. Higher values of α and β signify that CPT investors are becoming increasingly rational being. The parameter λ represents the loss aversion: the higher the value of λ, the more loss-averse the investor.

(2) The linear piecewise value function

$$v(t) = \begin{cases} t & t \geq 0 \\ \lambda t & t < 0 \end{cases} \tag{3.4}$$

[1] $\overline{\mathbb{R}}^+$ denotes $\mathbb{R}^+ \bigcup \{+\infty\}$ and $\overline{\mathbb{R}}^-$ denotes $\mathbb{R}^- \bigcup \{-\infty\}$.

This is a special case of (3) in which $\alpha = \beta = 1$. We call this the linear value function.

CPT investors do not weight outcomes by using objective probabilities. Besides, there exist different parameters in the weighing functions with the domains of gains and losses. The weighting functions are denoted by $w^+(\cdot)$ for gains and $w^-(\cdot)$ for losses. Tversky and Kahneman (1992) proposed the following functions:

$$w^+(F_D(t)) = \frac{F_D^\gamma(t)}{[F_D^\gamma(t) + (1 - F_D(t))^\gamma]^{1/\gamma}} \tag{3.5}$$

$$w^-(F_D(t)) = \frac{F_D^\delta(t)}{[F_D^\delta(t) + (1 - F_D(t))^\delta]^{1/\delta}} \tag{3.6}$$

where $F_D(t)$ is the cumulative distribution function of D, and $w^+ : [0,1] \to [0,1]$, $w^- : [0,1] \to [0,1]$ are non-decreasing and differentiable with $w^+(0) = w^-(0) = 0$ and $w^+(1) = w^-(1) = 1$. We call these the general weighting functions.

The parameters of the weighting functions determine the degree of distortion to the objective probabilities. The smaller the values of γ, δ, the greater the degree of distortion.

A number of scholars have focused on the comparative performance of different functional forms of CPT, examining a variety of different parameters relative to Tversky and Kahneman (1992). Camerer and Ho (1994) fitted the CPT function using three different value functions together with the general weighting function, whereas Wu and Gonzalez (1996) repeated their estimation procedures to determine the CPT parameters. Blondel (2002), Birnbaum and Chavez (1997) examined various functions and tested several parametric formulae under CPT. Fennema and Van Assen (1998) re-examined the utility for losses and derived parameters for the CPT value function. In addition, various experiments have examined different forms and parameters of functions under CPT Gonzalez and Wu (1999), Bleichrodt and Pinto (2000), Luce (2001). Table 3.1 presents a variety of CPT parameters.

It is interesting to discuss the value function as a piecewise linear function under CPT. Luce (1991) explored a weighted linear utility representation for binary gambles in which the weights depend on both the rank order and the sign of the consequences. Benartzi and Thaler (1995) gave a novel explanation for the equity premium puzzle using CPT with a linear value function. Barberis et al. (2001) systematically studied the asset prices given by adopting a linear value function based on CPT, and explained the high mean, excess volatility, and predictability of stock returns. Gruene and Semmler (2005)

Table 3.1: CPT parameters

Function	Parameter Value	Study
Value Function	$\alpha = 0.36$, $\beta = 0.24$	Fennema and Van Assen (1999)
	$\alpha = \beta = 0.50$	Wu and Gonzalez (1996)
	$\alpha = \beta = 0.88$	Tversky and Kahneman (1992)
Weighting Function	$\gamma = \delta = 0.56$	Camerer and Ho (1994)
	$\gamma = 0.61$, $\delta = 0.69$	Tversky and Kahneman (1992)

studied asset pricing by employing a stochastic growth model with linear function, and De Giorgi et al. (2007) developed an algorithm to compute asset allocations under CPT with a linear value function, thus extending the explanation of the equity premium puzzle by incorporating changing risk aversion. Schmidt and Zank (2008) and He and Zhou (2011) have also studied the linear value function under CPT.

The objective function of CPT investors is defined by[2]

$$\begin{aligned} V(D) = &-\int_0^{+\infty} v^+(t)dw^+(1-F_D(t)) \\ &+\int_{-\infty}^{0} v^-(t)dw^-(F_D(t)) \end{aligned} \tag{3.7}$$

or, equivalently,

$$\begin{aligned} V(D) = &\int_0^{+\infty} v^+(t)w^{+'}(1-F_D(t))f_D(t)dt \\ &+\int_{-\infty}^{0} v^-(t)w^{-'}(F_D(t))f_D(t)dt \end{aligned} \tag{3.8}$$

where $f_D(t)$ is the probability density function of D and

$$\begin{aligned} &w^{+'}(1-F_D(t)) \\ = \;& \frac{\gamma(1-F_D(t))^{\gamma-1}((1-F_D(t))^{\gamma}+(F_D(t))^{\gamma})}{[F_D^{\gamma}(t)+(1-F_D(t))^{\gamma}]^{1/\gamma}} \\ - \;& \frac{(1-F_D(t))^{\gamma}((1-F_D(t))^{\gamma-1}-(F_D(t))^{\gamma-1})}{[F_D^{\gamma}(t)+(1-F_D(t))^{\gamma}]^{1/\gamma}} \end{aligned} \tag{3.9}$$

[2] We use the definition of He and Zhou (2011).

$$
\begin{aligned}
& w^{-'}(F_D(t)) \\
&= \frac{\delta(F_D(t))^{\delta-1}((1-F_D(t))^{\delta}+(F_D(t))^{\delta})}{[F_D^{\delta}(t)+(1-F_D(t))^{\delta}]^{1/\delta}} \\
&- \frac{F_D(t)^{\delta}((F_D(t))^{\delta-1}-(1-F_D(t))^{\delta-1})}{[F_D^{\delta}(t)+(1-F_D(t))^{\delta}]^{1/\delta}}
\end{aligned} \tag{3.10}
$$

To ensure that (8) is finite, we require $\alpha < 2min(\gamma, \delta)$ and $\beta < 2min(\gamma, \delta)$. However, Barberis and Huang (2008) showed that these conditions are not necessary for log-normal and normal distributions.

For the convenience of computing, we give another form of the objective function.

Proposition 1: If the value function is the piecewise power function, the objective function of CPT can also be written as:

$$
\begin{aligned}
V(D) = & \int_0^{+\infty} w^+(1-F_D(t))dv^+(t) \\
& - \int_{-\infty}^{0} w^-(F_D(t))dv^-(t)
\end{aligned} \tag{3.11}
$$

Refer to Appendix A for the proof.

Assumption 2: There are no transaction costs in the financial market, and CPT investors do not borrow cash for investment.

Assumption 3: The financial market consists of n risky assets, whose rates of returns $\mathbf{R} = (\mathbf{R}_1, \mathbf{R}_2, \ldots, \mathbf{R}_n)^T$ follow the multivariate normal distribution, and one riskless asset with a return of r_0. Furthermore, the vector of means $E(\boldsymbol{R}) = \boldsymbol{\mu} = (\mu_1, \mu_2, \ldots, \mu_n)^T$ and the $n \times n$ covariance matrix $Cov(\mathbf{R}) = \Sigma = \{\sigma_{i,j}\}$ exists, where Σ is a positive-definite matrix.

Let $\boldsymbol{x} = (x_1, \ldots, x_n)$ be a vector of the investment ratio on risky assets, $x_{n+1} = (1 - \sum_{i=1}^{n} x_i)$ be the proportion of the riskless asset and $\mathbf{1}$ be the unit column vector. Then, the deviation $D(\boldsymbol{x})$ is

$$
\begin{aligned}
D(\boldsymbol{x}) &= x_1\mathbf{R}_1 + \cdots + x_n\mathbf{R}_n + x_{n+1}r_0 - r_f \\
&= \boldsymbol{x}\boldsymbol{R} - r_0\boldsymbol{x}\mathbf{1} - r_f + r_0
\end{aligned} \tag{3.12}
$$

In this dissertation, we only consider the case where $r_f = r_0$. That is, CPT investors take the return on the riskless asset as the reference point.

Proposition 2: If $\mathbf{R} \sim N_n(\boldsymbol{\mu}, \Sigma)$ holds, then the deviation $D(\boldsymbol{x}) \sim N(\mu_D, \Sigma_D^2)$, where

$$\begin{aligned} \mu_D &= \boldsymbol{x}\boldsymbol{\mu} - r_0\boldsymbol{x}\mathbf{1} \\ \sigma_D^2 &= \boldsymbol{x}\Sigma\boldsymbol{x} \end{aligned} \tag{3.13}$$

See Appendix for the proof.

CPT investors make portfolio choices by maximizing their CPT value; that is, CPT investors determine their investments by maximizing the value of equation (3.11)[3].

$$\begin{aligned} \max \quad & V(D(\boldsymbol{x})) \\ s.t. \quad & \sum_{i=1}^{n} x_i \leq 1, \\ & x_i \geq 0, i = 1, 2, 3 \ldots, n \end{aligned} \tag{3.14}$$

3.2 Adaptive real-coded genetic algorithm technique

The concept of genetic algorithm, which can be described as "intelligent" probabilistic search algorithms, was developed by Holland and his colleagues in the 1960s and 1970s. The idea was inspired by the evolutionary theory explaining the origin of species. Being a population-based approach, GAs are well suited to the CPT optimization. The ability of genetic algorithm to simultaneously search different regions of a solution space makes it possible to find a diverse set of solutions for difficulties with non-convex, discontinuous, and multi-modal solution spaces. Each of GAs initially has a population consisting of a set of vector chromosomes, which are generated randomly to explore the solution space of a problem.

Traditionally, the genes in the chromosome are represented by binary coded strings. RCGAs, however, use real-valued genes to solve continuous optimization problems (Herrera et al., 1998, Michalewicz, 1994). The use of real-valued genes allows for better adaptation to the numerical optimization of continuous problems. RCGAs have the capacity to exploit the gradation of functions with continuous variables, and to avoid

[3]The objective function (3.11), which involves numerical integration, is solved using Matlab's "integral" and "normcdf"functions.

the Hamming cliff effect suffered by BCGAs. The convergence speed of RCGAs is good, because, unlike BCGAs, where are no coding and decoding processes. In RCGAs, each chromosome represents one decision vector and every gene corresponds to the weight of one asset.

We developed an efficient RCGA with an adaptive mechanism that makes the operators more efficient throughout the evolutionary process. This adaptive method solves the problem of portfolio choice under CPT within the feasible operating region. There are two important issues in RCGA. One is the selection pressure, without which the search process would be a random algorithm. The effective selection pressure ensures that chromosomes with higher fitness values have a higher chance of surviving under crossover and mutation. The second is population diversity, which produces genotypes of the offspring that differ from those of their parents. A highly diverse population can increase the probability of exploring the global optimum and prevent premature convergence to a local optimum (Deb and Goyal, 1996, Deb and Beyer, 1999). RCGA involves a trade-off between selection pressure and population diversity, because the two factors act against one another. They should therefore be controlled to ensure the optimal balance. Hence, in this dissertation, we propose a new technique to increase the selection pressure and an adaptive method which retains the balance between the selection pressure and population diversity processes. The pseudocode of ARCGA is shown in Figure 3.1, where $P(g)$ represents the parents, $M(g)$ represents the mating pool, $Q'(g)$ represents the offspring from $M(g)$ after the crossover operation, $Q(g)$ represents the offspring from $Q'(g)$ after the mutation operation, and g denotes the generation. A flowchart of ARCGA is given in Figure 3.2.

3.2.1 Adaptive real-coded genetic algorithm implementation

Generation of initial population

The genes of a chromosome are real numbers between 0 and 1 representing the weights invested in the assets under CPT. The most popularly used initialization method is random generation. Every datum is generated uniformly in the range [0, 1] in a random, independent manner. For convenience, we add a variable x_{n+1} to ARCGA and $\sum_{i=1}^{n+1} x_i = 1$. The w_i, $i = 1, \ldots, n+1$, represent a vector of data generated randomly in the initialization phase.

If the sum of these data is greater than 1, the constraint in (15) will be violated. To

Begin
$g = 1$
Initialize population $P(g)$.
while (not termination condition) do
Evaluate the fitness value of each chromosome.
Sort the chromosomes in descending order of fitness value.
Calculate the adaptive parameters based on fitness value.
Create a mating pool $M(g)$ from the population $P(g)$ using the selection operator.
Create offspring-crossed $Q'(g)$ from $M(g)$ using the crossover operator.
Create offspring $Q(g)$ from $Q'(g)$ using the mutation operator.
Select the best chromosome from $P(g)$ according to fitness value and replace one chromosome at random in $Q(g)$.
$P(g+1) =$ normalized $Q(g)$.
$g = g + 1$
end
End

Figure 3.1: Pseudocode of the ARCGA

overcome this problem, the portfolio weights are obtained by normalizing w_i as follows:

$$x_i = w_i / \sum_{i=1}^{n+1} w_i \tag{3.15}$$

where x_i represents the weight invested in asset i after normalization, and x_{n+1} represents the proportion for the riskless asset.

By repeating the above procedure m times, we obtain m solutions that form the first population $X^{g=1} = \{X_1, \ldots, X_m\}$. X^g will evolve and gradually converge to X^* as the evolutionary process continues.

Constraints

One of the most difficult things in ARCGA is how to handle constraints. Dealing with infeasible chromosomes is far from trivial, because the genetic operators used to manipulate the chromosomes often yield infeasible solutions. Various methods have been proposed to deal with constraints, with the penalty method the most common approach for constrained optimization problems. The principal issue of the penalty strategy is how to design a penalty function so as to effectively guide the genetic search toward a promising area of the solution space. There are no general guidelines for designing a penalty function, and constructing an efficient penalty function is largely dependent on the specific problem. Actually, one reason for using the penalty method is to retain some information about some infeasible solutions, which are perhaps closer

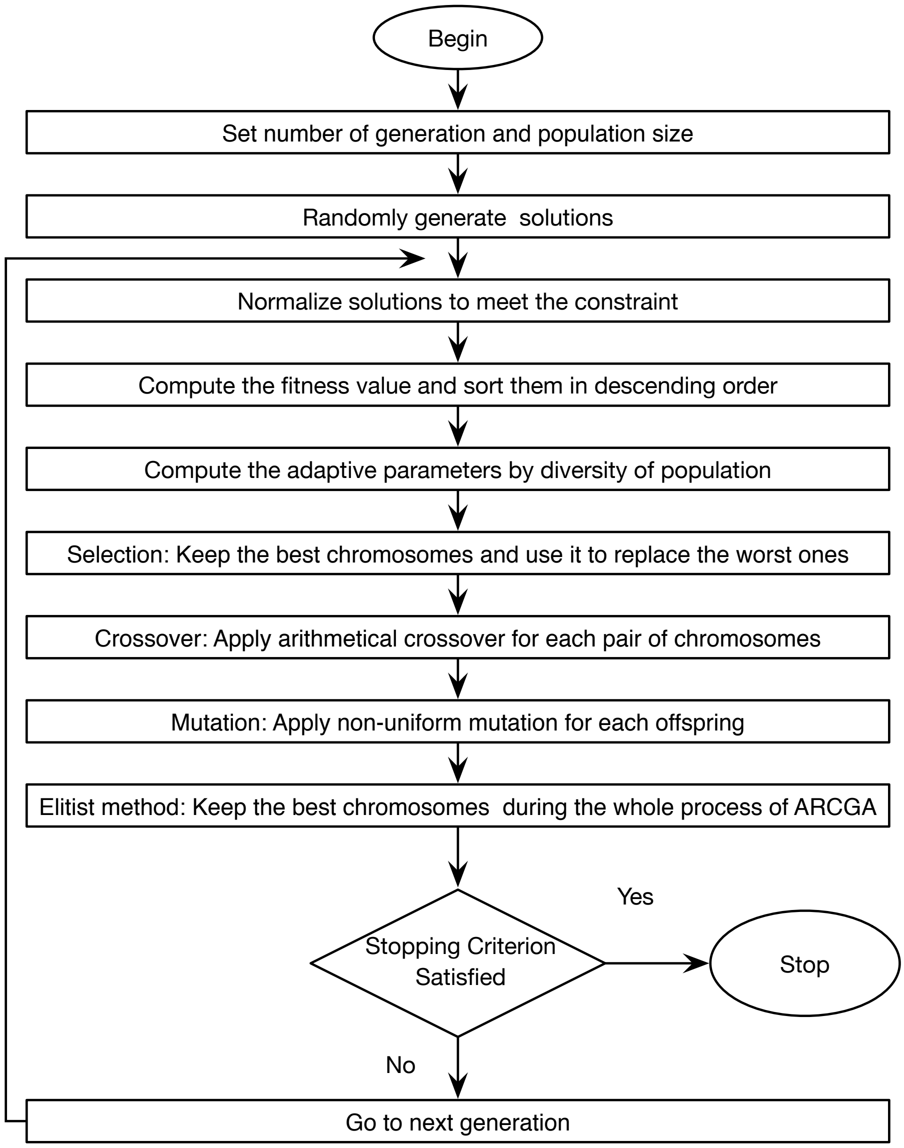

Figure 3.2: Flowchart of ARCGA

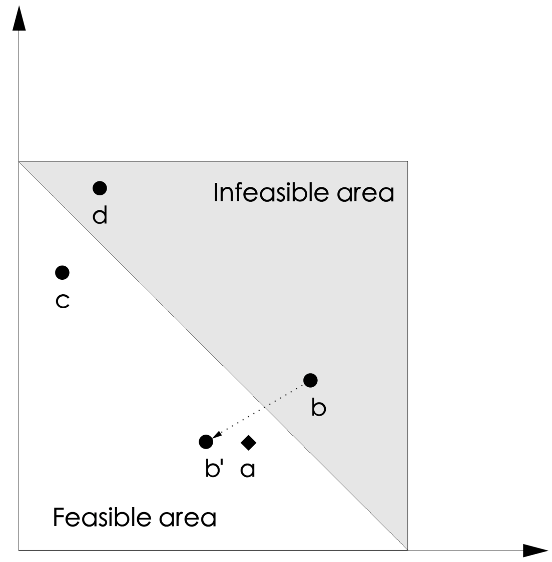

Figure 3.3: Feasible and infeasible solutions

to the global optimal solution. However, because of the complexity of CPT, it is very difficult to construct an effective penalty function.

To overcome this problem, we propose a system in which the normalization process is run after each iteration in the ARCGA process. As shown in Figure 3.3, in two dimensions, there exist two parts of the solution space: feasible and infeasible solutions. Infeasible solution b is much nearer to optimal solution a than infeasible solution d and feasible solution c. There are reasons to think that b contains much more information about the optimal solution than c, although it is infeasible. Solution b', with more information, is drawn into the feasible region by normalization.

Evaluation

The fitness value of each chromosome is measured by an objective function. We evaluate the fitness values of the chromosomes in $P(g)$ with (15), and order f_j^g, $j = 1, \ldots, m$, which are the fitness values in every generation.

Selection

There are different methods of applying the selection operator in RCGA. Truncation selection is known to be the most efficient form of directional selection (Crow and Kimura, 1979). This approach ranks all chromosomes according to their fitness values and selects the best $T\%$ as parents. Truncation selection has been used extensively in evolution strategies (Thierens and Goldberg, 1994a, Back, 1994). Other popular methods include

$(\mu + \lambda)$ selection and (μ, λ) selection, where μ is the number of parents and λ is the number of offspring. The top μ individuals form the next generation, with the selection being from parents and children in the $(\mu + \lambda)$ case and from children only for (μ, λ). Typically, λ is one or two times μ (Hoffmeister and Bäck, 1991). Top-N selection is employed to select the N best chromosomes from the population (Hancock, 1994). In addition, the "replace worst" strategy replaces the population if the new chromosome is better than the existing worst chromosome. Goldberg and Deb (1991) showed that higher selection pressure exists in populations that delete the worst chromosome, even if others are selected at random.

Inspired by these ideas, we introduce a selection operator called duplicated top-N selection (DTNS), in which the best n chromosomes are copied twice to the mating pool. This approach ensures that the best n chromosomes are retained and the worst n chromosomes are replaced. The remaining chromosomes are placed in the mating pool unchanged. In this manner, the best chromosomes in the population have more opportunity to be chosen and the worst chromosomes will be eliminated from the population. This leads to better convergence in terms of the quality of chromosomes and computation time. In our work, N is determined by rounding the size of the population and multiplying by T%.

To compare the validity of this technique with traditional methods, we use roulette wheel selection, which was proposed by Holland (1975) and is commonly used in GAs (Goldberg, 1989). Roulette wheel selection chooses chromosomes with respect to their fitness values (fitness-proportionate selection). Traditionally, fitness values are assigned to each chromosome based on their objective function values. In this work, we use the fitness value ranking, instead of the actual values, known as rank-based selection. This notion was first used in GAs by Baker (1985). There is some evidence that selection according to rank is superior to fitness-proportionate selection (Whitley et al., 1989). The use of rank-based roulette wheel selection (RRWS) provides a degree of control over the selective pressure that is not possible with fitness-proportionate roulette wheel selection. In addition, fitness-proportionate roulette wheel selection can sometimes lead to problems when the search is likely to stagnate due to a lack of selective pressure or premature convergence because selection has narrowed the search too quickly. Moreover, in some ways, ranking is more consistent with the schema theorem, because there is no need to introduce additional parameters that are not explained by the schema theorem in order to control the selective pressure. To some extent, the method of rank-based selection can prevent premature convergence, which is a weakness of RCGA.

As we are seeking to maximize the objective function, each chromosome will be sorted in descending order of objective function value, and the rank-based number used as the fitness value. Linear and exponential ranking methods are commonly used. We use the linear ranking and selection probabilities given by

$$p_i = (\eta^+ - (\eta^+ - \eta^-) \cdot (i-1)/(m-1))/m \tag{3.16}$$

where $\sum_{i=1}^{m} p_i = 1$ and $1 \le \eta^+ \le 2$, $\eta^- = 2 - \eta^+$.

The constants η^+ and η^- are called the maximum and minimum expected values, respectively, and determine the slope of the linear function. Normally, a value of $\eta^+ = 1.1$ is recommended Back (1994).

Crossover

We use the typical arithmetical crossover of each parent to produce two offspring in the crossover step. It was suggested by Michalewicz (1994) that the arithmetical crossover operator is the best option for RCGA. Assume that chromosomes $x = (x_1, \ldots, x_{n+1})$ and $x' = (x'_1, \ldots, x'_{n+1})$ have been selected for crossover. The offspring are given as follows:

$$\begin{aligned} \hat{x}_i &= \xi x_i + (1-\xi)x'_i; \\ \hat{x}'_i &= \xi x'_i + (1-\xi)x_i. \end{aligned} \tag{3.17}$$

where ξ is an uniform random number in $[-0.5, 1.5]$.

Mutation

Mutation is generally applied at the gene level and reintroduces genetic diversity to the population, which helps the search to escape from local optima. We apply a non-uniform mutation operator. Suppose that $x = (x_1, \ldots, x_i, \ldots, x_{n+1})$ is a chromosome, and that $x_i \in [a_i, b_i]$, where a_i and b_i are the lower and upper bounds of x_i, respectively, is the element to be mutated in generation g. The resulting chromosome will be $x' = (x_1, \ldots, x'_i, \ldots, x_{n+1})$, where x'_i is obtained by

$$x'_i = \begin{cases} x_i + \triangle(g, b_i - x_i) & \text{if } \gamma = 0 \\ x_i - \triangle(g, x_i - a_i) & \text{if } \gamma = 1 \end{cases} \tag{3.18}$$

with γ being a random number that takes a value of zero or one, and

$$\triangle(g, l) = l(1 - r^{(1-\frac{g}{G})^\tau}) \tag{3.19}$$

where r is a random number from the interval [0, 1], G is the maximal generation number, and τ is a user-selected parameter that determines the degree of non-uniformity (Michalewicz, 1994). This function gives a value in the range $[0, l]$ such that the probability of returning a number close to zero increases as g increases. As a result, this operator performs a uniform search in the initial stages (when g is small) and a more local search in the final stages (Kaelo and Ali, 2007, Deep and Thakur, 2007).

Elitist method

The main disadvantage of roulette wheel selection is that the best chromosome in each generation may not be preserved. The elitist method can isolate the best chromosome and transfer it to the next generation (Thierens and Goldberg, 1994b). In this way, the best chromosome obtained during the whole process of RCGA is guaranteed to survive. Rudolph (1994) showed that convergence to the global optimum is not an inherent property of the canonical genetic algorithm (CGA), but rather is a consequence of the algorithmic trick of keeping track of the best solution found over time. The major drawback of this approach is the tendency to get stuck around some local extrema (Ranković et al., 2014). Obviously, the combination of elitist and adaptive methods can avoid the aforementioned shortcomings.

Stop criterion

The algorithm terminates when the following stopping criterion is satisfied:

$$g = G \tag{3.20}$$

where g is the current number of generations and G denotes the maximum number of generations, which is a pre-fixed threshold.

3.2.2 Parameter selection

Although RCGA has many advantages over BCGA, it can often suffer from premature convergence due to a lack of population diversity. Conversely, it can also suffer from slow convergence (Kaelo and Ali, 2007). To overcome these problems, RCGA has been hybridized with other optimization methods (Chelouah and Siarry, 2003, Yun et al., 2003). For instance, some researchers have worked on improving the crossover operators of RCGA (Tsutsui and Goldberg, 2001, Hrstka and Kučerová, 2004). Population diversity in RCGA is important throughout the search process, not just in the initial stages, as this determines how the set evolves with each generation to explore the search region. Although a number of rules have been suggested in the literature to improve the population diversity, they are generally tailored towards solving certain problems (Tsutsui and Goldberg, 2001). Subbaraj et al. (2009) and Subbaraj et al. (2011) developed a self-adaptive real-coded genetic algorithm to solve the combined heat and power economic dispatch problem. However, p_c and p_m were assigned constant values in their paper.

In essence, RCGA uses the values of p_c and p_m to balance the capacity to converge to an optimum (local or global) after locating the region containing the optimum and the capacity to explore new regions of the solution space in search of the global optimum (Srinivas and Patnaik, 1994). The probabilities of crossover and mutation are varied depending on the fitness values of the solutions. This encourages the exploration of the search space because of the accelerating gene disruption, and helps to prevent premature convergence. Hence, in our study, we use the adaptive probabilities of crossover and mutation to maintain diversity in the population and sustain the convergence capacity of the RCGA. In general, GAs commonly use values of p_c in the range [0.5, 1] and p_m in the range [0.001, 0.05].

Srinivas and Patnaik (1994) used $f^g_{max} - \overline{f}$ to detect the convergence of the GA, and varied p_c and p_m depending on the value of this metric. Scholars are increasingly using adaptive crossover and mutation probabilities instead of fixed values (Lin et al., 2003, Blum et al., 2001, Zhang et al., 2007, Huimei, 2000). Following Srinivas and Patnaik (1994) and Lin et al. (2003), we use the following expressions:

$$
\begin{aligned}
p^g_c &= k_1(f^g_{max} - \overline{\overline{f}}^g)/(f^g_{max} - \overline{f}^g) \\
p^g_m &= k_2(f^g_{max} - \overline{\overline{f}}^g)/(f^g_{max} - \overline{f}^g)
\end{aligned}
\tag{3.21}
$$

We set $k_1 = 1.0$ and $k_2 = 0.5$. f^g_{max} and $\overline{f}^g$ represent the maximum and average fitness values of the population at each generation, respectively. $\overline{\overline{f}}^g$ is the average of those fitness values that are greater than $\overline{f}^g$. We restricted p_c and p_m to the ranges recommended above. The adverse effect caused by poor chromosomes can be avoided by calculating the difference in fitness values, and this is clarified in the degree of con-

Table 3.2: Parameters of objective function

Parameter	α	β	λ	γ	δ
Value	0.88	0.88	2.55	0.61	0.69

vergence between the chromosomes with larger fitness values in the population.

3.3 Numerical experiments

3.3.1 Parameters related to investors and futures markets

We test the proposed algorithm for solving portfolio selection model (5.21). To perform computational experiments, we first determine the parameters in model (5.21). Five parameters $(\alpha, \beta, \lambda, \gamma, \delta)$ are used to describe the investors' objective function. We adopt the values from Tversky and Kahneman (1992). Table 3.2 lists the parameters used in our experiments.

In this study, we consider that the multivariate normal distribution of risky assets has mean vector and covariance matrix as follows:[4]

$$\mu = \begin{pmatrix} 0.040 & -0.015 & 0.039 & 0.027 \end{pmatrix} \tag{3.22}$$

$$\Sigma = 10^{-5} * \begin{pmatrix} 11.45 & 4.69 & 2.29 & 7.01 \\ 4.69 & 16.90 & 3.16 & 6.02 \\ 2.29 & 3.16 & 12.77 & 2.96 \\ 7.01 & 6.02 & 2.96 & 24.34 \end{pmatrix} \tag{3.23}$$

The return on the riskless asset is set to 0.02.

3.3.2 Computational experiments

All algorithms are programmed in MATLAB R2014b and run on a 2.6 GHz Apple MacBook Pro Computer with 8 GB RAM.

We first compare the method of DTNS with RRWS. The number of iterations and population size are set as 150 and 100 tentatively. Figures 3.4 and 3.5 show the statistic for the number of iteration size from 0 to 150. The dash-dot curve in Figure 3.4 represents the maximum of 100 fitness values per iteration .The dotted curve in Figure 3.4 represents the average of 100 fitness values per iteration. The dash-dot curve in

[4] The data are derived by Hu and Kercheval (2010) and the four stocks are Disney, Pfizer, Altria, Intel.

Figure 3.5 represents the maximum of 100 fitness values per iteration. The dotted curve in Figure 3.5 represents the average of 100 fitness values per iteration. The horizontal axis is the number of iterations and the vertical axis is the fitness value. As shown in 3.4 and 3.5, other things being equal, for the time taken to find the optimal solution or the trend of convergence, the DTNS method we proposed is significantly superior to RRWS. And, 150 iterations should suffice with our next experiments.

The population size is very important in ARCGA, which significantly influences the convergence of the algorithm. We set the number of iterations to 150 and test population sizes of 30, 50 respectively. Figures 3.6 and 3.7 show the statistic for the number of iteration size from 0 to 150. The dash-dot curve in Figure 3.6 represents the maximum of 50 fitness values per iteration, while the dotted curve in Figure 3.6 represents the average of 50 fitness values per iteration. However, the dash-dot curve in Figure 3.7 represents the maximum of 30 fitness values per iteration. But the dotted curve in Figure 3.7 represents the average of 30 fitness values per iteration. The horizontal axis is the number of iterations and the vertical axis is the fitness value. However, as shown in the Figures 3.5,3.6 and 3.7, we observe no significant differences in finding the optimal solutions and the trend of convergence.

We then compare the computer time under population size 30,50 and 100. For each given population size, the ARCGA is run 100 times. We find that the results obtained by the 100 runs are the same for different population sizes. In addition, we use an exhaustive method to find the global solutions to three decimal places. A comparison shows that the results given by ARCGA and the exhaustive method are exactly the same. The computational results show that the proposed ARCGA is a rapid, effective, and stable algorithm for the given objective function. The results are summarized in Table 3.3. As shown in Table 3.3, the computer time of population size at 30 is significantly less than others. For the experiments in the next section, we set the population size at 30.

Table 3.3: Computational results

PS	CV	FR	x_1^*	x_2^*	x_3^*	x_4^*	x_5^*	$Time^a$
30	0.0271	0.0396	0.6261	0	0.3739	0	0	24
50	0.0271	0.0396	0.6261	0	0.3739	0	0	40
100	0.0271	0.0396	0.6261	0	0.3739	0	0	76

Note: PS stands for Population size; CV stands for CPT Value; FR stands for Final Return; x_1^*, x_2^*, x_3^*, x_4^* and x_5^* stand for optimal investment ratios for stock Disney, Pfizer, Altria, Intel and riskless asset respectively.

[a] Average time in seconds.

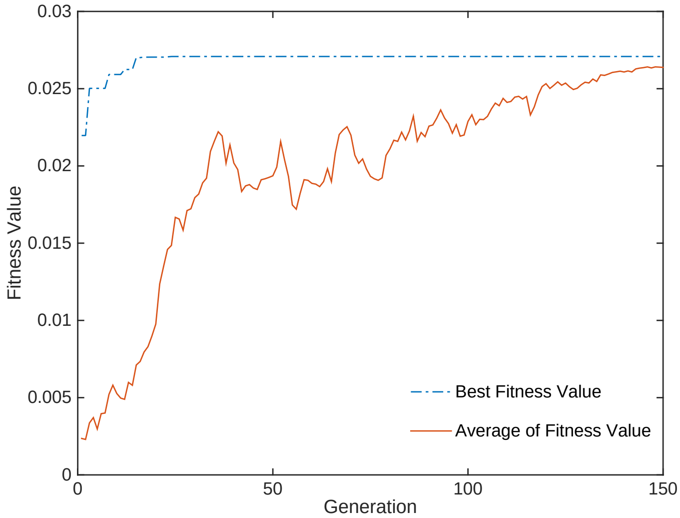

Figure 3.4: Fitness values produced by RRWS with population size 100.

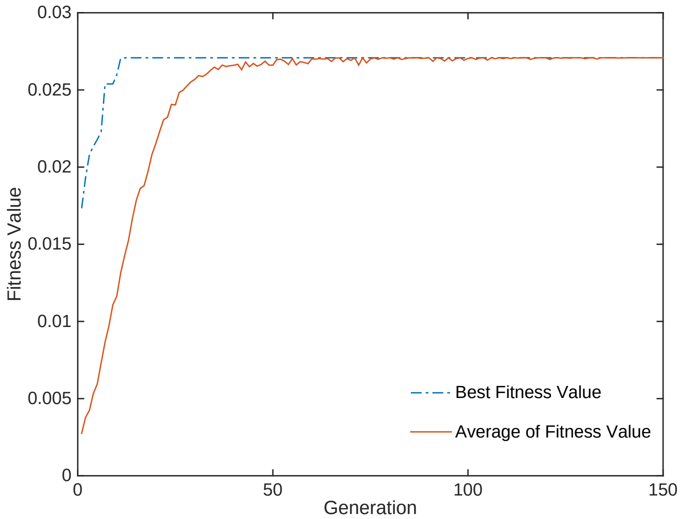

Figure 3.5: Fitness values produced by DTNS with population size 100.

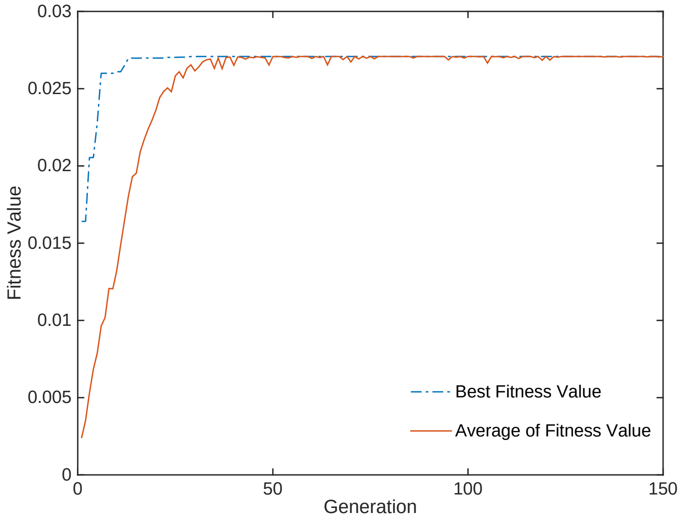

Figure 3.6: Fitness values produced by DTNS with population size 50.

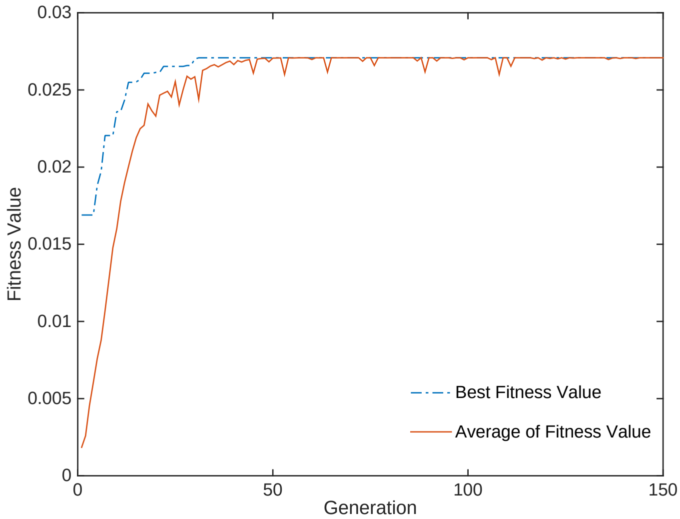

Figure 3.7: Fitness values produced by DTNS with population size 30.

3.4 Influence of parameters on objective function

We now examine how the objective function is influenced by different parameter values. Although CPT has gained importance in recent years, there has been little research on the portfolio choice problem because of the complexity of the CPT function, and the influence of the different parameters it is largely unknown. Pirvu and Schulze (2012) studied how an investor allocated portfolio in a continuous model under CPT. However, they only studied the influence of parameters α and γ on CPT. Coelho et al. (2014) studied the parameters of CPT in a discrete optimization model of portfolio selection. Asset allocation, however, has not been studied in his paper. The proposed ARCGA allows us to analyze the portfolios given by different parameters. As mentioned before, each set of parameters listed in Table 1 represents the preference of different CPT investors. We considered three values of the loss aversion parameter λ: 1.55 to represent a low degree of loss aversion; 2.55 to represent a normal loss aversion, as discussed by Tversky and Kahneman (1992); and 3.55 to represent a high degree of loss aversion.

Tables 3.4–3.7 report the CPT values, final returns, and optimal portfolio for every asset, as well as the average runtimes.[5] The algorithm converged to the same result for each set of parameters with population sizes of 30, 50, and 100, which again proves that ARCGA is a stable algorithm. The results indicate that, regardless of the parameter values, most CPT investors invest all their money in the risky Disney and Altria stocks. One explanation for this is that the close to riskless returns are unattractive for these investors, to say nothing of the risky asset with negative yield. It is clear that the CPT values decrease with an increase in loss aversion, other conditions being equal, which is consistent with our intuition. Furthermore, we can see that the difference in the final return is not obvious under different parameters.

It can be observed from Table 3.4 that, along with the increasing values of α, β, the CPT values have declined rapidly. However, the CPT values increase slightly with γ and δ. CPT investors will increase the proportion of Disney stock under the conditions mentioned above. In addition, our results suggest that α and β are more sensitive than γ and δ to the CPT values.

Tables 3.5–3.7 indicate that there is at least one linear case among the value and weighting functions. From these tables, it can be observed that the loss aversion λ has little effect on the CPT values. As shown in Table 3.7, CPT investors will invest all

[5] All results in these tables are global optimal solutions given by the exhaustive method to three decimal places.

Table 3.4: Results for solution of CPT with general value function and general weighting function

	λ	γ	δ	CV	FR	x_1^*	x_2^*	x_3^*	x_4^*	x_5^*	Time[a]
α=0.36	1.55	0.56	0.56	0.1712	0.0396	0.5599	0	0.4401	0	0	117
β=0.24	2.55	0.56	0.56	0.1532	0.0396	0.5573	0	0.4427	0	0	122
	3.55	0.56	0.56	0.1352	0.0396	0.5562	0	0.4438	0	0	126
	1.55	0.61	0.69	0.1926	0.0396	0.5628	0	0.4372	0	0	112
	2.55	0.61	0.69	0.1825	0.0396	0.5600	0	0.4400	0	0	118
	3.55	0.61	0.69	0.1723	0.0396	0.5584	0	0.4416	0	0	119
α=0.50	1.55	0.56	0.56	0.1075	0.0396	0.5700	0	0.4300	0	0	21
β=0.50	2.55	0.56	0.56	0.1030	0.0396	0.5652	0	0.4348	0	0	21
	3.55	0.56	0.56	0.0985	0.0396	0.5623	0	0.4377	0	0	21
	1.55	0.61	0.69	0.1162	0.0396	0.5770	0	0.4230	0	0	22
	2.55	0.61	0.69	0.1138	0.0396	0.5714	0	0.4286	0	0	23
	3.55	0.61	0.69	0.1114	0.0396	0.5678	0	0.4322	0	0	21
α=0.88	1.55	0.56	0.56	0.0257	0.0396	0.6178	0	0.3822	0	0	26
β=0.88	2.55	0.56	0.56	0.0250	0.0396	0.5959	0	0.4041	0	0	26
	3.55	0.56	0.56	0.0244	0.0396	0.5846	0	0.4154	0	0	25
	1.55	0.61	0.69	0.0274	0.0397	0.6590	0	0.3410	0	0	24
	2.55	0.61	0.69	0.0271	0.0396	0.6261	0	0.3739	0	0	24
	3.55	0.61	0.69	0.0268	0.0396	0.6088	0	0.3912	0	0	23

Note: For each set of parameters, the results are the same over the 100 executions under population sizes of 30, 50, and 100.
CV stands for CPT Value; FR stands for Final Return; x_1^*, x_2^*, x_3^*, x_4^* and x_5^* stand for optimal investment ratios for stock Disney, Pfizer, Altria, Intel and riskless asset respectively.
[a] Average time in seconds under population size 30.

Table 3.5: Results for solution of CPT with linear value function and general weighting function

	λ	γ	δ	CV	FR	x_1^*	x_2^*	x_3^*	x_4^*	x_5^*	Time[a]
α=1	1.55	0.56	0.56	0.0165	0.0397	0.6669	0	0.3331	0	0	13
β=1	2.55	0.56	0.56	0.0162	0.0396	0.6201	0	0.3799	0	0	14
	3.55	0.56	0.56	0.0158	0.0396	0.6002	0	0.3998	0	0	13
	1.55	0.61	0.69	0.0176	0.0398	0.7958	0	0.2042	0	0	13
	2.55	0.61	0.69	0.0174	0.0397	0.6831	0	0.3169	0	0	13
	3.55	0.61	0.69	0.0172	0.0396	0.6446	0	0.3554	0	0	13

Note: For each set of parameters, the results are the same over the 100 executions under population sizes of 30, 50, and 100.
CV stands for CPT Value; FR stands for Final Return; x_1^*, x_2^*, x_3^*, x_4^* and x_5^* stand for optimal investment ratios for stock Disney, Pfizer, Altria, Intel and riskless asset respectively.
[a] Average time in seconds under population size 30.

Table 3.6: Results for solution of CPT with general value function and linear weighting function

	α	β	λ	CV	FR	x_1^*	x_2^*	x_3^*	x_4^*	x_5^*	Time[a]
γ=1	0.36	0.24	1.55	0.2316	0.0396	0.5892	0	0.4108	0	0	89
δ=1	0.36	0.24	2.55	0.2292	0.0396	0.5790	0	0.4210	0	0	100
	0.36	0.24	3.55	0.2268	0.0396	0.5738	0	0.4262	0	0	103
	0.50	0.50	1.55	0.1350	0.0396	0.6366	0	0.3634	0	0	13
	0.50	0.50	2.55	0.1345	0.0396	0.6202	0	0.3798	0	0	13
	0.50	0.50	3.55	0.1340	0.0396	0.6091	0	0.3909	0	0	12
	0.88	0.88	1.55	0.0312	0.0400	0.9577	0	0.0423	0	0	23
	0.88	0.88	2.55	0.0311	0.0398	0.8348	0	0.1652	0	0	23
	0.88	0.88	3.55	0.0310	0.0398	0.7767	0	0.2233	0	0	23

Note: For each set of parameters, the results are the same over the 100 executions under population sizes of 30, 50, and 100.
CV stands for CPT Value; FR stands for Final Return; x_1^*, x_2^*, x_3^*, x_4^* and x_5^* stand for optimal investment ratios for stock Disney, Pfizer, Altria, Intel and riskless asset respectively.
[a] Average time in seconds under population size 30.

Table 3.7: Results for solution of CPT with linear value function and linear weighting function

		λ	CV	FR	x_1^*	x_2^*	x_3^*	x_4^*	x_5^*	Time[a]
α=1	γ=1	1.55	0.0199	0.0400	1	0	0	0	0	12
β=1	δ=1	2.55	0.0198	0.0400	1	0	0	0	0	13
		3.55	0.0197	0.0399	0.9162	0	0.0838	0	0	13

Note: For each set of parameters, the results are the same over the 100 executions under population sizes of 30, 50, and 100.
CV stands for CPT Value; FR stands for Final Return; x_1^*, x_2^*, x_3^*, x_4^* and x_5^* stand for optimal investment ratios for stock Disney, Pfizer, Altria, Intel and riskless asset respectively.
[a] Average time in seconds under population size 30.

their money in Disney stock unless the loss aversion exceeds a certain threshold.

In general, not all assets are selected by CPT investors. In addition, although there are major differences amongst the CPT values, different parameters are not so influential on the final returns. Unlike EUT, which has only one parameter, there are several parameters with different characteristics in CPT, and their influence on portfolio choice varies. Moreover, different investment behavior might occur in the stock market under some extreme cases.

3.5 Further discussion on CPT under normal distribution

It requires enormous computing for determining a portfolio that maximizes the expected utility of final wealth in the next period. Some scholars have demonstrated that a point on the efficient frontier curve can yield maximum expected utility when investor's utility function is quadratic, or the probability distribution of returns is joint elliptical(Levy and Markowitz, 1979, Meyer, 1987, Kroll et al., 1984, Ross, 2011). Markowitz (2014) suggested that it is much more convenient and economical to determine a set of mean-variance efficient portfolios than it is to fine the portfolio which maximizes expected utility.

Based on the Taylor series expansion around return in the next period，an investor's utility function may be expanded as follows:

$$\begin{aligned} u(R) &= u(E(R)) + u'E(R)(R - E(R)) + \frac{1}{2}u''E(R)(R - E(R))^2 \\ &+ \sum_{n=3}^{\infty} \frac{1}{n!} u^{(n)} E(R)(R - E(R))^n \end{aligned} \tag{3.24}$$

Take the expected value of both sides of the equation (3.24), the investor's expected utility can be expressed as

$$E(u(R)) = u(E(R)) + \frac{1}{2}u''E(R)(Var(R))^2 + \sum_{n=3}^{\infty}\frac{1}{n!}u^{(n)}E(R)\nu^n(R) \tag{3.25}$$

where $\nu^n(R)$ denotes the n-th central moment of R.

It can be proved that equation (3.25) has the property of being both increasing in $E(R)$ and decreasing in $Var(R)$ if $u(\cdot)$ is a nondecreasing and concave function and R obeys a normal distribution (Ross, 2011).

Levy and Levy (2004) discussed the relation between prospect theory and mean-variance rule and employ the mean-variance optimization algorithm to construct efficient portfolios of prospect theory. And, they employed stochastic dominance rules to prove that the PT and mean-variance efficient sets almost coincide. That is the set of making the CPT function reach maximizing is a subset of efficient frontier. However, they did not study the relationship between value of PT and expected return or its variance. As much as we know, no paper has so far been discussed about relationship between value of PT or CPT and expected return or its variance. The problem mentioned above can be solved in this dissertation.

It has known that the CPT value is function of both portfolio's mean and its variance according to equation (3.11).

Proposition 3: Under Assumption 1 and 2, if the portfolio return follows a normal distribution, the value function takes the form in equation (3.2) and the probability weighting functions take the form in equation (3.5) and (3.6) with $\gamma = \delta = 1$. We have the following conclusions:

(1) CPT value is an increased function of mean of portfolio return on the whole interval.

(2) CPT value is an increased function of variance of portfolio return in negative interval. CPT value can be either a non-increased or non-decreased function of variance of portfolio return in positive interval. Consequently, it needs to be further investigation that whether the CPT value is a non-increased function or non-decreased function of variance of portfolio return on whole interval.

3.6 Summary

There has recently been increased research interest in CPT, but most studies have not taken portfolio optimization into account. In this study, we built a single-period portfolio selection model for CPT investors in a market consisting of one riskless asset and several risky assets.

Considering the complex nature of CPT, which is a non-convex, non-concave, and non-smooth function, we proposed a real-coded genetic algorithm with adaptive operators to solve the model.

Computational experiments were conducted to demonstrate that the proposed AR-CGA outperformed traditional GAs. All results were the same after 100 executions with different population sizes, which shows that the approach proposed in this dissertation efficiently and effectively solves the problem of portfolio choice under CPT.

We have also presented the first study of the influence of various parameters on the CPT values. The experimental results showed that α and β are more sensitive than γ and δ. Not all assets in the portfolio were selected by the CPT investors, and different investment behavior might occur in the stock market under certain extreme cases.

And moreover, we studied the relationship between CPT value and the mean of portfolio return and the standard deviation of portfolio return under multivariate normal distribution.

These cases provide a reference for further research into the portfolio choice problem under CPT.

Chapter 4

Portfolio choice under scenarios

In this chapter, a method of coupling scenario techniques for simulating the scenario of the real stock market with a genetic algorithm to determine the optimal solution is presented. The major challenge is to provide data on mathematical models in determining optimal solutions to address uncertainties in the field of financial investment.

Due to the effectiveness of the mathematical models hinges on the quality of the scenarios. Bradley and Crane (1972) first employed and presented these techniques to the financial world using several scenario generation methods support financial decision making. This chapter therefore focuses on three different bootstrap method to achieve scenario generation.

The bootstrap method is a way of resampling in statistics, and was first introduced by Efron (1992). The key idea is to provide a resampling simulation technique to estimate the complicated characteristics of the underlying population. The bootstrap method does not generate random variates, but repeatedly samples the original data (Efron and Tibshirani, 1993) instead. It is a highly effective tool in the absence of a parametric distribution for obtaining a set of data. The bootstrap method is thus used especially when the number of available samples is relatively small and a larger number of observations is required. In fact, the use of bootstrapping for scenario generation has been suggested by Kouwenberg and Zenios (2006). Generally, in the analysis of financial time series, the probability distribution of a data set is unknown; the bootstrap method is suitable for assessing the distribution properties of some statistic of such data.

As mentioned in the previous chapter, genetic algorithms (GAs), as the robust search and optimizations techniques, have been successfully used in various fields. In recent years, numerous studies have shown that GAs can help to solve optimal portfolio problems in finance. However, to our knowledge, no studies have incorporated GAs with

Bootstrap method to solve the portfolio choice problem under CPT.

This chapter presents a CPT model for optimal portfolio selection under scenario and couples the bootstrap method for the evaluation of investment portfolio scenarios with genetic algorithm to determine the optimal solution. Sufficient empirical comparisons of different choices under different reference points in the CPT model are provided in this chapter.

4.1 Objective function for CPT investors

Under **Assumption 1-2**, consider a set of investment assets $i = 1, 2, \ldots, n$. At the end of a certain holding period these assets generate random returns $\mathbf{R} = (R_1, R_2, \ldots, R_n)^T$. The CPT investors attempt to allocate their budget to these assets by deciding on a specific allocation $\mathbf{x} = (x_1, x_2, \ldots, x_n)^T$, $x_i \geq 0$ (no short sales permitted) and $\sum_{i=1}^{n} x_i = 1$ (basic budget constraint). Using the vector $\mathbf{1} = (1, 1, \ldots, 1)^T$, we may write the basic budget constraints in vector form as

$$X = \{\mathbf{x} : \mathbf{x}^T \mathbf{1} = 1, \mathbf{x} \geq 0\}$$

At the next horizon period the uncertain return of the portfolio is denoted by $R_p = \mathbf{x}^T \mathbf{R} = \sum\limits_{i=1}^{n} x_i R_i$. This indicates that the current selection may impact on future investment returns.

Let r_f be the value of a (scalar) reference point that separates gains and losses, we define the deviation Y from the reference level by

$$Y = R_p - r_f \tag{4.1}$$

Obviously, Y is treated as a random variable. Suppose that $Y_1, \ldots, Y_m$ form a random sample with some distribution. Let Y_1 denote the smallest value in the random sample, Y_2 denote the next smallest value, and so on. In this way, Y_m denotes the largest value in the sample, and Y_{m-1} denotes the next largest value. Thus, the random variables $Y_1, \ldots, Y_m$ are the order statistics of the sample. Let $y_1, \ldots, y_i, y_0, y_{i+1}, \ldots, y_m$ denote the values of the order statistics for an arbitrarily sample with probability $p_1, \ldots, p_i, p_0, p_{i+1}, \ldots, p_m$, respectively. If all values of $y_1, \ldots, y_m$ are nonzero, then $y_0 = 0$ with probability $p_0 = 0$ is inserted; otherwise, there exists $y_0 = 0$ with probability $p_0 \neq 0$. (In fact, the presence or absence of this zero value in the results has no effect on the CPT value, as will be seen later.) According to Tversky and Kahneman Tversky and Kahneman (1992), the CPT investors are supposed to evaluate the investment as the following:

$$(y_1, p_1; \ldots; y_i, p_i; y_0, p_0; y_{i+1}, p_{i+1}; \ldots; y_m, p_m) \tag{4.2}$$

The original version of the PT suffers from potential violations of first-order stochastic dominance, and PT can be applied only to gambles with at most two nonzero outcomes. We therefore apply CPT by Tversky and Kahneman (1992). When the probabilities are weighted in PT, however, it is the cumulative probabilities that are weighted in CPT:

$$\pi_i = \begin{cases} \pi_i^+ = w^+(p_i + \cdots + p_m) - w^+(p_{i+1} + \cdots + p_m) \\ \pi_i^- = w^-(p_1 + \cdots + p_i) - w^-(p_1 + \cdots + p_{i-1}) \end{cases} \tag{4.3}$$

where i denotes outcome y_i $(i = 1, \ldots, m)$.

By using the equation (3.3) as value function and use equations (3.5) and (3.6) as probability weighting functions, we present the CPT value of the investment for stocks by

$$V(x) = \sum_{i=1}^{m} \pi_i \cdot v(y_i(x)) \tag{4.4}$$

CPT investors tend to make portfolio choices by maximizing their CPT value; that is, CPT investors determine their investments by maximizing the value of equation (4.4). We now formally propose the following objective function:

$$\begin{array}{ll} \max & V(x) \\ s.t. & \sum\limits_{i=1}^{n} x_i = 1, \\ & x_i \geq 0, i = 1, \ldots, n \end{array} \quad . \tag{4.5}$$

4.2 Bootstrap method

4.2.1 Non-parametric method

A critical problem in portfolio selection is the description of a random investment portfolio return, and the problem is generally addressed by a set of random returns or their expected return. Thus a set of scenarios can be generated by different methods, such as a historical approach, bootstrap method, or Monte Carlo simulation. In this chapter, the past observations of asset returns are used to generate the expected returns by using the bootstrap method. Specifically, we need to combine historical data with the bootstrap technique to simulate the required number of sample data. However, it is difficult to determine the parameters of the returns. Thus, the non-parametric bootstrap (NPB) method is used in this chapter.

Bootstrap method simulates what would happen if we sample repeatedly from the basic set and also records observing data that the available data are created through available data, i.e. resampling (Härdle et al., 2012). The dimension of the created data is smaller than the dimension of the original data, achieving the best results if the bootstrap random selections are designed with replications and having the same dimension as original data (Franke et al., 2004).

As shown in the Figure 4.1, new random data are drawn by the way of sampling with replacement，with equal size original set. As we observed that the statistical analysis can be obtained by using of B sampling. Figure 4.1 describes the basic principle of this method. Bootstrapping repeatedly draws random samples, each with size n, from the original sample of size n, and makes the probability 1/ n for each original variate in each sampling with replacement.

r		r_1^*	r_2^*	$\cdots$	r_B^*
r_1		r_2	r_4	$\cdots$	r_5
r_2	random sampling with replacement	r_4	r_1	$\cdots$	r_4
r_3	$\longrightarrow$	r_3	r_3	$\cdots$	r_3
r_4		r_4	r_3	$\cdots$	r_4
r_5		r_1	r_5	$\cdots$	r_5

Figure 4.1: Basic principle of the bootstrap method.

Suppose $R_1, R_2, \ldots, R_n$ be independent, identically distributed real random variables with a distribution function F. Let $\theta = \theta(F)$ be an unknown parameter of F which needs to be estimated. Although there are many statistical estimations, this chapter focuses only on the expected value. The one big advantage of non-parameter bootstrap method is that the distribution of data does not need to be known in advance. The unknown distribution F can be replaced by the empirical distribution function as shown in equation (4.6) function using of bootstrap method.

$$\hat{F}_n(\chi) = \frac{1}{n}\sum_{i=1}^{n} \mathbb{1}_{\{R_i \le \chi\}} \tag{4.6}$$

where $\mathbb{1}_{\{A\}}$ is the indicator of event A and χ denotes the number of elements being less than or equal to χ in the sample.

4.2.2 Bootstrap methods for financial time series

Consider a strictly stationary time series of the ith investment asset held for T time periods. This is expressed by $\mathbf{R}_i = (R_{i,1}, \ldots, R_{i,T})$, which means that the joint probability distribution of $(R_{i,1}, \ldots, R_{i,T})$ does not change when shifted in time. As mentioned earlier, it is difficult to find the probability distribution of $\mathbf{R}_i$, which is denoted by F_i. Let $\theta(F_i)$ be some parameter of interest such as the mean, median, or standard deviation of F_i. Let $\hat{\theta}(\mathbf{R}_i)$ be an estimator of $\theta(F_i)$ computed using observations $\mathbf{R}_i$. Here, we focus on the mean of the returns.

The bootstrap method does not require any parametric assumption on F_i, but can utilize smaller sample sizes as a formalization of the resampling procedure for statistical inference. Let the observed data take the values $\mathbf{r}_i = (r_{i,1}, \ldots, r_{i,T})$. The mean return of one asset is

$$\bar{r}_i = \frac{1}{T}\sum_{j=1}^{T} r_{i,j} \tag{4.7}$$

and the expected portfolio return at time index $T+1$ is expressed as

$$r_p = \sum_{i=1}^{n} x_i \bar{r}_i \tag{4.8}$$

Then, draw T sample data $\mathbf{r}_i^* = (r_{i,1}^*, \ldots, r_{i,T}^*)$ from $(r_{i,1}, \ldots, r_{i,T})$ by using of bootstrap method. The mean

$$\bar{r}_i^* = \frac{1}{T}\sum_{i=1}^{T} r_{i,T}^* \tag{4.9}$$

can be computed from $\mathbf{r}_i^*$. Note that the number of sampled data in the bootstrap method is equal to the number of observed data, and there is no permutation because we have performed random sampling without replacement.

By repeating this process S times, we obtain the scenario matrix:

$$\mathbf{R}_s = \begin{bmatrix} \bar{r}_1^{*1} & \bar{r}_1^{*2} & \cdots & \bar{r}_1^{*S} \\ \bar{r}_2^{*1} & \bar{r}_2^{*2} & \cdots & \bar{r}_2^{*S} \\ \vdots & \vdots & \ddots & \vdots \\ \bar{r}_n^{*1} & \bar{r}_n^{*2} & \cdots & \bar{r}_n^{*S} \end{bmatrix} \tag{4.10}$$

A vector $\mathbf{R}_p^S = \mathbf{x}^T \mathbf{R}_s = (r_p^1, \ldots, r_p^s)$ is obtained from the scenario matrix through multiplication by a set of decision-making vectors. If the elements in $\mathbf{R}_p^S$ are sorted in ascending order of value, we obtain result similar to (4.2). A set of optimal decision-making vectors that maximize the objective function will be discussed later.

Generally, selecting the best bootstrap technique for estimating the mean depends on whether the observed data $\mathbf{R}_i$ are assumed to be independent or dependent, which can be difficult to identify. Scenario generation, however, should be considered to encompass all future possibilities. We will refer to the bootstrap method for independent data as the standard bootstrap (SB) technique and that for dependent data as the moving block bootstrap (MBB) technique and non-overlapping block (NBB) technique.

Standard bootstrap

The SB method is implemented by sampling the data randomly with replacement,i.e., observed data can be resampled with a constant probability $1/T$. We can derive $\mathbf{R}_i^* = (R_{i,1}^*, \ldots, R_{i,T}^*)$ from $(R_{i,1}, \ldots, R_{i,T})$. For a more comprehensive review of the SB technique, see Ref.Kouwenberg and Zenios (2006).

Moving block bootstrap methods

Note that $\mathbf{R}_i = (R_{i,1}, \ldots, R_{i,T})$ is treated as a series of outcomes with probability $1/T$. However, this assumption is not always valid, especially for financial time series. Singh (1981) showed that SB technique, as considered in Hall (1985) for independent data, failed to produce valid approximations in the presence of dependent data. To overcome the limitations of the SB technique for dependent financial time series data, Hall Hall (1985) suggested resampling the data using blocks of observed data instead of individual data, and Kunsch (1989) advocated resampling blocks of observations at a time (see also Bühlmann and Künsch (1995)). Besides, the dependence structure of the random variables at short lag distances is preserved by keeping the neighboring observations together within the blocks. As a result, resampling blocks allows one to carry this information over to the bootstrap variables. Thus, a similar method was named the "moving block bootstrap" Liu and Singh (1992).

Suppose that $\mathbf{R}_i = (R_{i,1}, \ldots, R_{i,T})$ is the observed financial time series of the ith assets. Let ℓ be an integer satisfying $1 \leq \ell < T$. Define the overlapping blocks $\mathbf{B}_{i,1}, \ldots, \mathbf{B}_{i,M}$ of length ℓ as

$$\begin{aligned}
\mathbf{B}_{i,1} &= (R_{i,1}, \ldots, R_{i,\ell}) \\
\mathbf{B}_{i,2} &= \quad (R_{i,2}, \ldots, R_{i,\ell+1}) \\
&\ldots \\
\mathbf{B}_{i,M} &= \qquad (R_{i,T-\ell+1}, \ldots, R_{i,T})
\end{aligned} \tag{4.11}$$

where $M = T - \ell + 1$. To generate the MBB samples, we select $b = T/\ell$ blocks at random with replacement from $(\mathbf{B}_{i,1}, \mathbf{B}_{i,2}, \ldots, \mathbf{B}_{i,M})$. Because each resampled block has ℓ elements, concatenating the elements of the b resampled blocks serially yields $T = b \cdot \ell$ bootstrap observations. Some typical choices of ℓ are $\ell = CT^{1/k}$, for $k = 3, 4$, where $C \in R$ is a constant (Kreiss and Lahiri, 2012).

Non-overlapping block bootstrap

Another bootstrap technique involves resampling from non-overlapping blocks to generate the bootstrap observations Carlstein (1986). Suppose that ℓ is an integer in $[1, T]$ (note that NBB is equivalent to SB when ℓ=1). Let $N = T/\ell$ and generate NBB samples by selecting N blocks at random with replacement from the collection $(\bar{\mathbf{B}}_{i,1}, \bar{\mathbf{B}}_{i,2}, \ldots, \bar{\mathbf{B}}_{i,N})$ where

$$\begin{aligned}
\bar{\mathbf{B}}_{i,1} &= (R_{i,1}, \ldots, R_{i,\ell}) \\
\bar{\mathbf{B}}_{i,2} &= (R_{i,\ell+1}, \ldots, R_{i,2\ell}) \\
&\ldots \\
\bar{\mathbf{B}}_{i,N} &= (R_{i,(N-1)\ell+1}, \ldots, R_{i,T})
\end{aligned} \tag{4.12}$$

It is easier to examine the characteristics of the NBB estimators than those of the MBB estimators of a population parameter, because NBB uses non-overlapping blocks. However, the NBB estimators typically have higher MSEs for a given block size ℓ compared with their MBB counterparts Lahiri (1999).

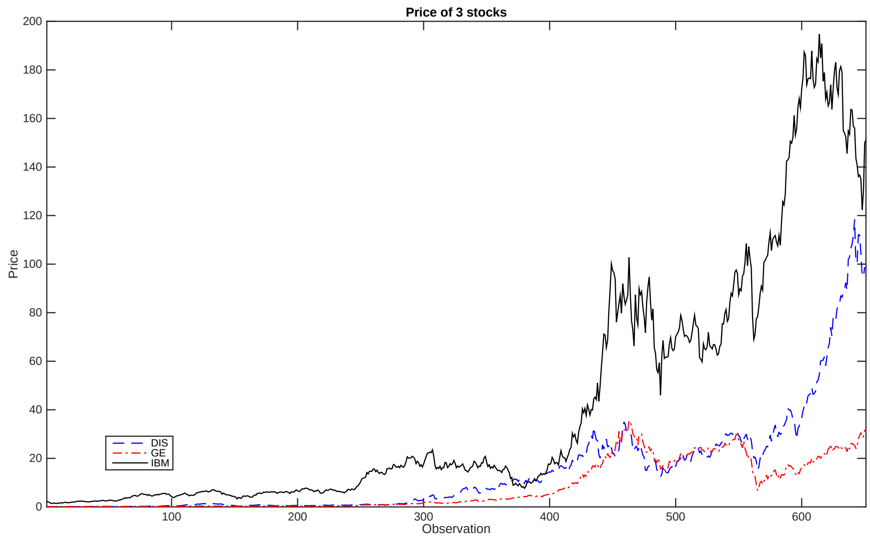

Figure 4.2: The prices of 3 stocks.

4.3 Numerical computation experiments

4.3.1 Parameters of the CPT investors and data

As mentioned in chapter 3, the CPT investors' objective function has five parameters α, β, γ, δ, and λ, which were proposed by Tversky and Kahneman (1992), as shown in Table 3.2.

Choosing the historical period is important for generating scenarios, but no rule for determining the length of the time period. Given the monthly return, we select a longer interval as soon as possible. In this dissertation, we consider portfolios composed of Walt Disney (DIS), General Electric (GE), and International Business Machines (IBM) stocks, and use the adjusted monthly closing price for the period January 2, 1962, to April 1, 2016 (dividends are not included). As much as 651 observations have been conducted in total, as shown in Figure 4.2 (data taken from Yahoo Finance).

The difference between the log and arithmetic returns is negligible for one-day horizon. However, some typical errors in the portfolio log-returns will be found if we neglect the conversion between the log-return and arithmetic return over much longer horizons. As a result, the difference between the log and arithmetic returns is generally taken into account, so we use the arithmetic returns and to obtain 650 monthly returns for each of the three stocks considered in this dissertation, as shown in Figure 4.3-4.5.

The foundation of time series analysis is stationarity. First, we test for the unit root of stock returns, which tells us whether a time series variable is non-stationary and

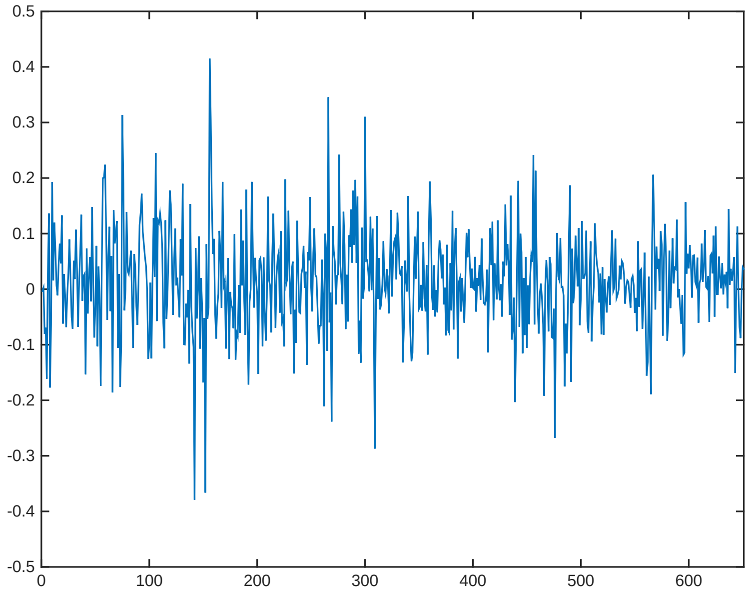

Figure 4.3: The monthly returns for DIS stock.

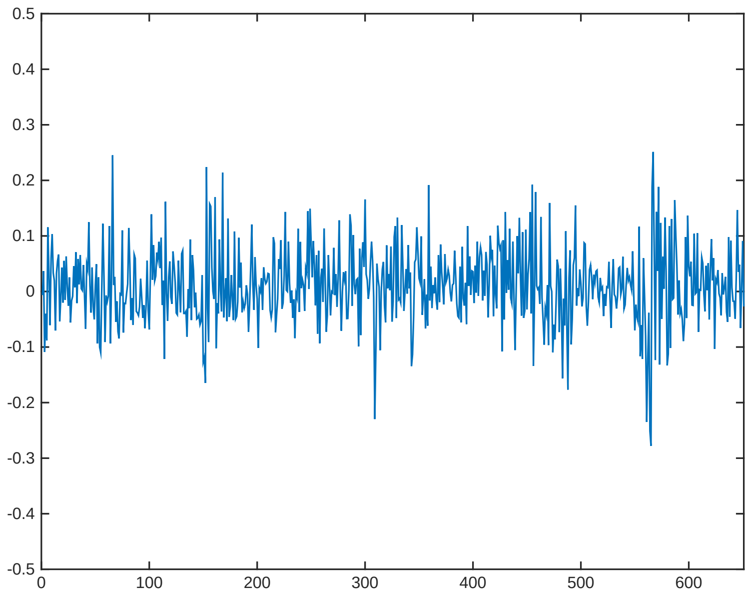

Figure 4.4: The monthly returns for GE stock.

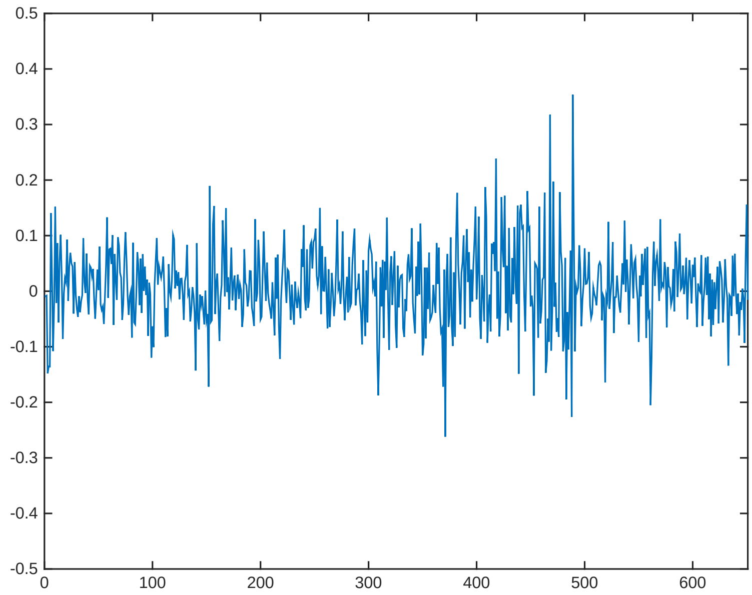

Figure 4.5: The monthly returns for GE stock.

Table 4.1: Descriptive statistics and test results for sample data

Company	Mean	Std.Dev	Skewness	Kurtosis	JB test	ADF test
DIS	0.0155	0.0895	0.0042	4.9735	1	0.001
GE	0.0106	0.0686	0.0540	4.2515	1	0.001
IBM	0.0089	0.0696	0.2270	4.8843	1	0.001

possesses a unit root. The results of ADF tests show that the returns on each of the three stocks reject the null hypothesis at the 1% significance level, i.e., there are no unit roots in the three sequences, and they can be considered as stationary sequences (see Table 4.1). We now analyze the descriptive statistics about the rate of return. Table 4.1 presents the basic statistics of the sample data. What we observed is that each of the stocks exhibits positive expected returns. Interestingly, the standard deviation of the GE returns is less than that of the IBM stocks, but GE's stock has a higher expected return than IBM's and is preferred by rational people as will be discussed later. Besides the IBM stock, the skewness is small, indicating that the distributions are largely symmetric. Each of the stocks exhibits bigger kurtosis values and has a heavy-tailed distribution. The study suggest the JB test shows that none of the stocks is normal at the 1% significance level. The normality can also be tested using a Q-Q plot. If the data is normally distributed, then the quantiles will lie on a straight line. Fig. 4.6-4.8 show the significant deviation from the straight line in the tails for each stock,

Table 4.2: Descriptive statistics and test results for simulation data

	Compay	Mean	Std.Dev.	Skewness	Kurtosis	JB test
SB	DIS	0.0154	0.0035	-0.0030	2.9830	0
	GE	0.0106	0.0027	0.0143	2.9401	0
	IBM	0.0089	0.0027	0.0317	2.9334	0
NBB	DIS	0.0155	0.0032	-0.0157	3.0165	0
	GE	0.0107	0.0026	-0.0120	2.9701	0
	IBM	0.0089	0.0025	-0.0011	2.9803	0
MBB	DIS	0.0160	0.0011	0.0207	2.9612	0
	GE	0.0107	0.0009	-0.0148	3.0711	0
	IBM	0.0093	0.0009	0.0517	2.9895	0

especially in the lower tail, indicating that the distribution of standardized returns is more heavy-tailed than the normal distribution.

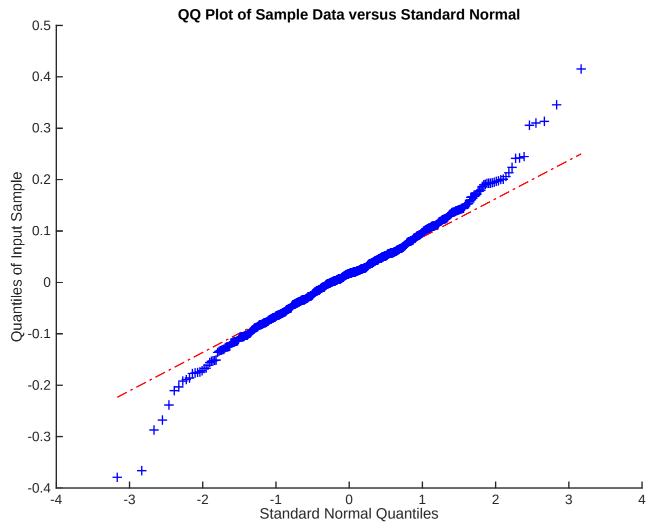

Figure 4.6: Q-Q plots for DIS

4.3.2 Computational experiments

All algorithms are programmed in MATLAB 2014b and are run on a 2.6GHz and 8GB RAM Apple MacBook Pro Computer.

First, we set the block length to 10 for MBB and NBB. Second, we generate 650 data at random with replacement using the SB technique and 65 blocks at random with replacement using MBB and NBB techniques. Third, we repeated these procedures 10,000 times to produce three scenario matrices, i.e., SB matrix, MBB matrix, and NBB

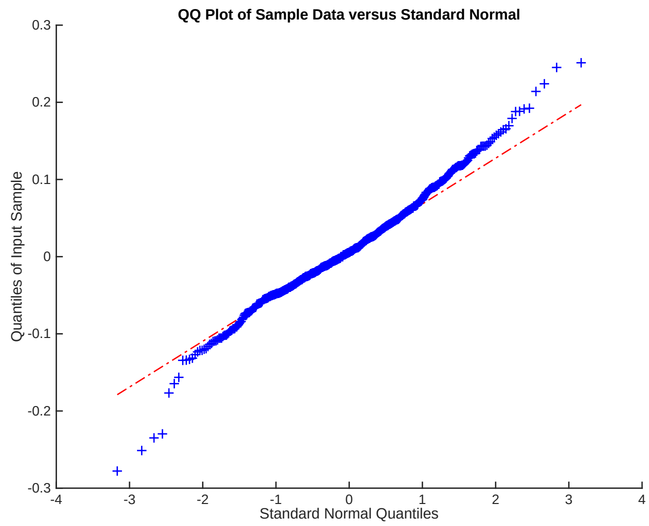

Figure 4.7: Q-Q plots for GE

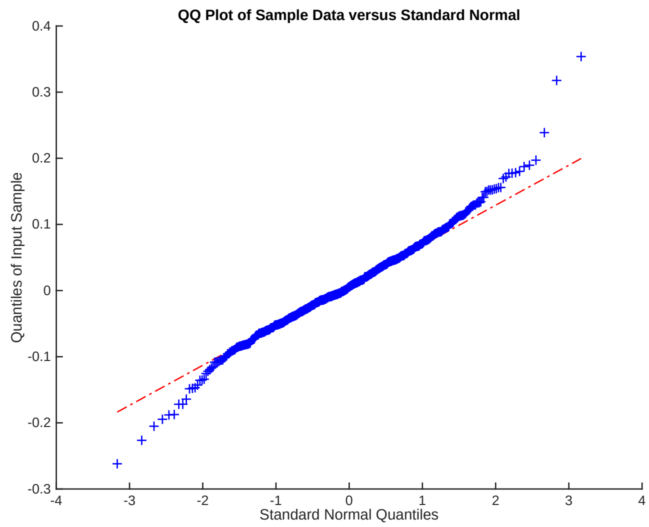

Figure 4.8: Q-Q plots for IBM

matrix, in which each row corresponds to one stock and each column represents every result of a single simulation. We can find the the relevant statistics for the simulation samples presented in Table 4.2. There is no significant change in the mean compared with the original data. However, the standard deviation, skewness, and kurtosis of the sample data for the three stocks have changed dramatically. All of the sample data passed the JB test for normality at the 1% significance level.

Generally, reference points are an important concept in CPT, but there has been little research into their impact on investment behaviors. Several different reference point scenarios have been discussed by Pirvu and Schulze (2012), providing no further details.

Table 4.3: CPT value, portfolio return and optimal solution with different reference point

	r_f	CPT VALUE	Return	DIS^*	GE^*	IBM^*
SB	0.003	0.0194	0.0154	1	0	0
	0.004	0.0165	0.0143	0.8047	0.0913	0.1040
	0.005	0.0135	0.0131	0.6050	0.1814	0.2136
	0.006	0.0080	0.0123	0.4442	0.3102	0.2455
NBB	0.003	0.0191	0.0151	0.9075	0.0925	0
	0.004	0.0169	0.0145	0.7914	0.2086	0
	0.005	0.0146	0.0139	0.6755	0.3245	0
	0.006	0.0114	0.0128	0.4708	0.4418	0.0874
MBB	0.003	0.0215	0.0160	1	0	0
	0.004	0.0200	0.0160	1	0	0
	0.005	0.0185	0.0160	1	0	0
	0.006	0.0169	0.0160	1	0	0

As mentioned above, each column of the scenario matrixes represents every result of a single simulation. We can obtain a CPT value by coupling a solution generated by ARCGA with one scenario matrix. By using continuously generating solutions, we identify the optimal solution that maximizes the objective function. The ARCGA was used to produce different initial populations of sizes 50, 100, and 150, and the algorithm was executed 100 times. All 100 results were the same for each population size. The results are presented in Table 4.3, where r_f denotes the reference point value, Return denotes the portfolio return, and the superscript $*$ denotes the optimal investment ratio for that stock.

The computational results show that the proposed ARCGA is an effective and stable algorithm for the given objective function. For comparison, we use an exhaustive method to find the global optima. The results show that these solutions are exactly equal to those generated by the proposed ARCGA.

Generally, the different scenarios have a significant influence on investment behaviors under CPT, and the MBB scenario has the most significant effect. Regardless of how the reference point changes, CPT investors always put all of their money into DIS under the MBB scenario. Our second observation is that CPT investors tend to change their investment ratio significantly under SB and NBB as the value of r_f changes. Furthermore, as r_f increases, the CPT values decrease significantly and the returns become smaller, which demonstrates numerically that greater expectations lead to greater disappointment.

4.4 Summary

In this chapter, we formulated an optimal portfolio selection model for a single period under CPT. Considering that the objective function is non-convex, non-concave, and non-smooth, we proposed an ARCGA to determine the optimal solution. We compared the portfolio choices of CPT investors based on different bootstrap techniques for scenario generation. Computational experiments show that the ARCGA can efficiently and stably solve the portfolio problem for CPT investors. In addition, this is the first study to consider the impact of different reference points on investment behavior under CPT.

Chapter 5

Portfolio choice with constraints

Currently, financial regulators propose some risk management requirements in terms of losses. Mathematically, risk management is a process of how to control the loss distributions. Value-at-risk (VaR) and conditional value-at-risk (CVaR) are popular tools for managing risk. This chapter solves the portfolio optimization under CPT with VaR constraints, CVaR constraints, and other relevant constraints. And, the optimal portfolios under various constraints are provided in this chapter.

5.1 Risk constraints

Suppose that $\boldsymbol{R} = (R_1, R_2, \ldots, R_n)^T$ is the return of each risky asset and $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)^T$ is the proportion in each risky asset. The loss of a portfolio over a fixed period can be defined as

$$L(\boldsymbol{x}, \boldsymbol{R}) = -\sum_{i=1}^{n} x_i R_i = -\boldsymbol{x}^T \boldsymbol{R} \tag{5.1}$$

For the normal distribution and discrete distribution, according to equation (2.18) the following VaR formulate can be easily to verify.

When the Loss L is normal distribution with mean μ and the standard variance σ, i.e $L \sim N(\mu, \sigma^2)$, then

$$VaR_c(\boldsymbol{x}, \boldsymbol{R}) = \mu + \sigma \Phi^{-1}(c) \tag{5.2}$$

where Φ represents the standard Normal distribution function, as shown Figure (5.1).

The loss L is a discrete distribution which takes the values $-\boldsymbol{x}^T \boldsymbol{R}^i$ with equal probability. The vectors $\boldsymbol{R}^i, i = 1, 2, \ldots, N$ are called the scenarios. VaR_c can be expressed as follow:

$$VaR_c(\boldsymbol{x}, \boldsymbol{R}) = M_{[\lfloor cN \rfloor : N]}(-\boldsymbol{x}^T \boldsymbol{R}^1, \ldots, -\boldsymbol{x}^T \boldsymbol{R}^N) \tag{5.3}$$

where $M_{[k:N]}(u^1, \dots, u^N)$ denotes the k-largest among $(u^1, \dots, u^N)$. It means that $M_{[1:N]}$ represents the minimum and $M_{[N:N]}$ denotes the maximum. The floor function $\lfloor \vartheta \rfloor$ is called the greatest integer function or integer value, gives the largest integer less than or equal to ϑ (Spanier and Oldham, 1987, Graham et al., 1994). As shown in Figure 5.2.

For discrete scenarios, VaR_c is a nonconvex, discontinuous function, which may lead to difficulties when computing optimal portfolios.

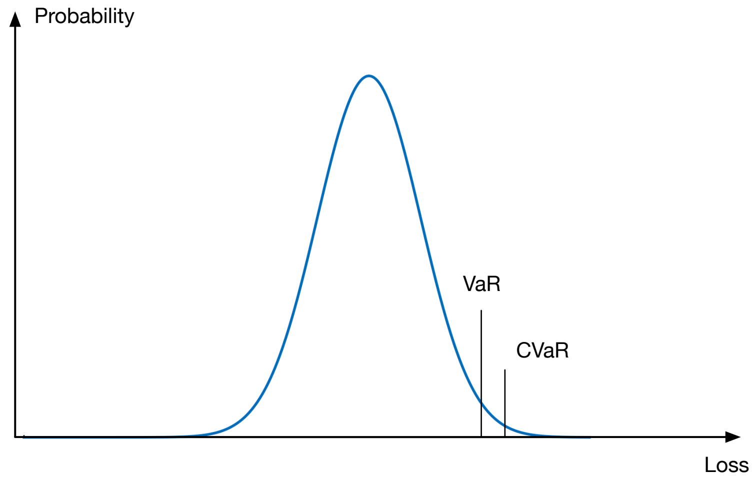

Figure 5.1: VaR and CVaR with normally distribution

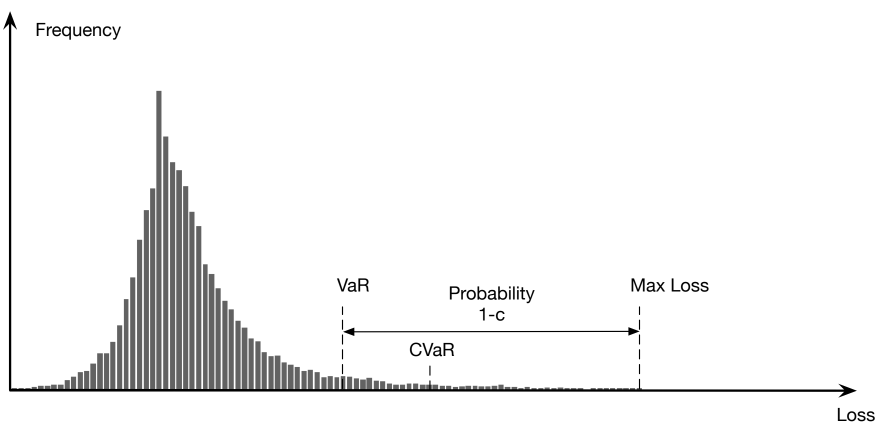

Figure 5.2: VaR and CVaR with discrete distribution

Generally, risk is considered as a correction factor for expected returns before the presentation of the MPT by Markowitz (1952). Markowitz proposed to measure the risk according to variance and semi-variance. However, it has been criticized that standard deviation cannot well explain the phenomenon of “fat tails” in financial investment field and penalizes ups and downs from the mean equally.

Risk management has received much attention from practitioners and regulators in the last few years, with VaR emerging as one of the most popular tools. Jorion (2006), Linsmeier and Pearson (2000), Alexander and Baptista (2002), Chance (2004) noted that VaR is widely used as a risk management tool by corporate treasurers, dealers, fund managers, financial institutions, and regulators, etc. To measure risk, we need to establish a relationship between the random variables, such as loss, and a non-negative real number, i.e., $\mathcal{R} : R \to \mathbb{R}$. The scalar measure of risk allows to sort and to compare investments based on their respective risk values. Artzner et al. (1999) introduced the concept of coherent risk measure and have extensively criticized the use of VaR as a measure of risk. Rockafellar (2007) proposed a functional $\mathcal{R} : \mathcal{L}^2 \to (-\infty, \infty)$ as a coherent risk measure in the extended sense if [1]

- (R1): $\mathcal{R}(C) = C$ for all constants C;
- (R2): $\mathcal{R}((1-\lambda)R + \lambda R') \leq (1-\lambda)\mathcal{R}(R) + \lambda\mathcal{R}(R')$ for $\lambda \in (0,1)$ (convexity);
- (R3): $\mathcal{R}(R) \leq \mathcal{R}(R')$ when $R \leq R'$ (monotonicity);
- (R4): $\mathcal{R}(R) \leq 0$ when $\| R^k - R \|_2 \to 0$ with $\mathcal{R}(R^k) \leq 0$ (closedness);
- (R5): $\mathcal{R}(\lambda R) = \lambda\mathcal{R}(R)$ for $\lambda > 0$ (positive homogeneity)

The property of subadditivity can be obtained from combination of (R1) with (R2).

$$\mathcal{R}(R + R') \leq \mathcal{R}(R) + \mathcal{R}(R') \tag{5.4}$$

Subadditivity means diversification that total risk of portfolio assets less than or equal to sum of risk of each asset.

VaR can summarize risk in a number, becoming the standard measure for financial analysts to quantify market risk. Many scholars have made use of VaR as constraint for problem of portfolio selection. Gaivoronski and Pflug (1999) solved a problem of maximizing portfolio with constraints of acceptable VaR. Yiu (2004) studied the optimal portfolio problem when a VaR constraint is imposed. They provided a way to control risks in the optimal portfolio and to full the requirement of regulators on market risks. Kleindorfer and Li (2005) solved the multi-period optimal portfolio problem under VaR

[1] $\mathcal{L}^p$ function space is defined as:

$$\|e\|_p = (\sum_{j=1}^{m} |e_j|^p)^{1/p}$$

with $p = 1$ representing the absolute measure (or Manhattan distance), $p = 2$ the standard deviation (or Euclidean distance) where we can use variance instead because of the monotonic transformation property, and $p = \infty$ represents the largest absolute value where we can represent the losses for Minimax optimization.

constrained and showed its relationship to efficient frontier analysis in standard portfolio theory.

VaR, although a popular tool, has been controversial because of mathematical shortcoming and it can not measure the risk of extreme events exceeding VaR. VaR has been criticized by some scholars. Mausser and Rosen (1999) showed that VaR can be ill-behaved as a function of portfolio positions and can exhibit multiple local extrema, which can be a major handicap in trying to determine an optimal mix of positions or even the VaR of a particular mix. Basak and Shapiro (2001) showed that when an agent faces a VaR constraint at the initial date in a continuous-time model, the agent may select a larger exposure to risky assets than he or she would have chosen in its absence. VaR is not a coherent measure of risk, unless the underlying distribution is elliptical, such as normal distribution, student's distribution etc. That is, VaR fails to satisfy the property of subadditivity and the VaR of a portfolio with two securities may be larger than the sum of the VaR of the securities in the portfolio. For these reasons, the aforementioned researchers have proposed using conditional value-at-risk (CVaR) rather than VaR. Risk often was restricted that can be assumed by a strict subset of all investors.

CVaR has proved to be superior than VaR in some respects (Rockafellar and Uryasev, 2000, 2002). CVaR, introduced by Rockafellar and Uryasev (2000), is a popular tool for managing risk as well. As an alternative measure of risk, CVaR is known to have better properties than VaR. CVaR, is the probability-weighted average of tail losses, or losses exceeding VaR (Dowd, 2003).

The CVaR can surpass the measure of VaR defined as the maximum loss at a specified confidence level which is commonly used as constraints in financial analysis in recent years. Alexander and Baptista (2004) analyze the portfolio selection implications arising from imposing a VaR constraint on the mean variance model, and compare them with those arising from the imposition of a conditional CVaR constraint. Boudt et al. (2013) used CVaR as budgets to analyze asset allocation. Krokhmal et al. (2002) solved the optimization problems for maximizing expected returns with CVaR constraints. Tian et al. (2010) extended approach proposed by Krokhmal et al. (2002) throughout adding CVaR-like constraints to the traditional portfolio optimization problem. Yamai and Yoshiba (2005) illustrated the way how the tail risk of VaR can cause serious problems in certain cases and CVaR can play its role.

For a continuous loss distribution, the CVaR at confidence level $c \in (0, 1)$ for loss L

is expressed as

$$
\begin{aligned}
CVaR_c(\boldsymbol{x}, \boldsymbol{R}) &= \frac{1}{1-c}\int_{VaR_c}^{+\infty} l\, dF_L(l) \\
&= \frac{1}{1-c}\int \mathbb{1}_{\{L \geq VaR_c\} l f_L(l) dl}
\end{aligned} \tag{5.5}
$$

If L is normal distribution $N(\mu, \sigma^2)$, then

$$
CVaR_c(\boldsymbol{x}, \boldsymbol{R}) = \mu + \sigma \frac{\varphi(\Phi^{-1}(c))}{1-c} \tag{5.6}
$$

where $\Phi(\cdot)$ is the standard normal distribution function and $\varphi(\cdot)$ is the standard normal density function, as shown in Figure 5.1.

For discrete scenarios, similar to equation (5.3) and as shown in the Figure 5.2, CVaR can be calculated by

$$
CVaR(\boldsymbol{x}, \boldsymbol{R}) = \frac{1}{N}\frac{1}{1-c}\sum_{-\boldsymbol{x}^T\boldsymbol{R}^i \geq VaR_c} -\boldsymbol{x}^T\boldsymbol{R}^i \tag{5.7}
$$

Actually, VaR and CVaR measure different parts of the distribution and have different properties of mathematics. For some companies, they tend to prefer VaR to CVaR that VaR may be far less than CVaR with the same confidence level. Thus they don't have to pay more for underlying losses, always preferring to provide reports to shareholder and regulators according to VaR. VaR may be better for optimization portfolio whereas CVaR may not perform well on certain conditions (Sarykalin et al., 2008).

As far as we know, no papers research in portfolio optimization under CPT coupled with constraints of VaR and CVaR, save for the Pirvu and Schulze (2012). These make it possible for us to comment on the differences presented in this dissertation and Pirvu and Schulze (2012). First, Pirvu and Schulze (2012) imposed VaR and CVaR as constraints for solving the ill-posedness problem, which is non-existence in our model. Second, Pirvu and Schulze (2012) did not consider how the constraints of VaR and CVaR will affect behavior of CPT investor, which is one of focuses in this dissertation. Several different situations about reference point are proposed in their paper, but they failed to provide further details.

Then the portfolio optimisation problem of CPT with a VaR risk constraint can be written as

$$
\begin{aligned}
\max \quad & V(D(\boldsymbol{x})) \\
s.t. \quad & VaR_c < l \\
& \sum_{i=1}^{n} x_i \leq 1, \\
& x_i \geq 0, i = 1, 2 \ldots, n
\end{aligned}
\tag{5.8}
$$

where $V(D(\boldsymbol{x}))$ is the equation (3.11).

The portfolio optimisation problem of CPT with a CVaR risk constraint can be written as

$$
\begin{aligned}
\max \quad & V(D(\boldsymbol{x})) \\
s.t. \quad & CVaR_c < l \\
& \sum_{i=1}^{n} x_i \leq 1, \\
& x_i \geq 0, i = 1, 2 \ldots, n
\end{aligned}
\tag{5.9}
$$

where $V(D(\boldsymbol{x}))$ is the equation (3.11).

There are other constraints such as portfolio return or investment proportions of the some assets except VaR being a constraint of investment for control the risk. To simplify the analysis，several problems are considered based on the objective function (3.11) as follows [2]:

$$
\begin{aligned}
\max \quad & V(D(\boldsymbol{x})) \\
s.t. \quad & \sum_{i=1}^{n} x_i \leq 1, \\
& \check{r}_0 < x_0 < \hat{r}_0 \\
& x_i \geq 0, i = 1, 2 \ldots, n
\end{aligned}
\tag{5.10}
$$

[2] x_0 is the investment proportion for risk-free asset, i.e. $x_0 = 1 - \sum_{i=1}^{n} x_i$. $(\check{r}_0, \hat{r}_0)$ is interval constraint of investment proportion for risk-free asset.

$$\begin{aligned}
\max \quad & V(D(\boldsymbol{x})) \\
s.t. \quad & \sum_{i=1}^{n} x_i \mu_i + x_0 r_f > \mathring{r} \\
& \sum_{i=1}^{n} x_i \leq 1, \\
& \check{r}_0 < x_0 < \hat{r}_0 \\
& x_i \geq 0, i = 1, 2 \ldots, n
\end{aligned} \tag{5.11}$$

$$\begin{aligned}
\max \quad & V(D(\boldsymbol{x})) \\
s.t. \quad & VaR_c < l \\
& \sum_{i=1}^{n} x_i \leq 1, \\
& \check{r}_0 < x_0 < \hat{r}_0 \\
& x_i \geq 0, i = 1, 2 \ldots, n
\end{aligned} \tag{5.12}$$

5.2 Deviation constraints

VaR is usually expressed as a positive number for the worst loss at given confidence level. VaR has implicit meaning that it is a relative concept (Jorion, 2006). The *relative* VaR is defined as the loss relative to the mean on the horizon:

$$VaR(\mu) = E(W) - W^* = -W_0(R^* - \mu) \tag{5.13}$$

where W_0 is the initial investment and R is its rate of return, $W^* = W_0(1 + R^*)$ is the lowest portfolio value at given confidence c.

Generally, VaR is defined as the *absolute* VaR which is relative to zero or without reference to the expected value:

$$VaR(0) = W_0 - W^* = -W_0 R^* \tag{5.14}$$

The mean return is very small under short investment horizons, where both relative VaR and absolute VaR may give similar results. Jorion (2006) argued that relative VaR is conceptually more appropriate because it regards risk according to a deviation from the mean on the target date, appropriately accounting for the time value of money. We can say that the method of relative VaR is more conservative when the mean value is positive and consistent with definitions of unexpected loss, becoming common for measuring credit risk over long horizons.

Generally, statistics are most often obtained from a set of data by simulations. The expected value of portfolio returns, for example, obtained by the bootstrap method shown in chapter 4, are always positive. It is difficult to work with absolute VaR, being not necessary to be considered as another measure.

Rockafellar et al. (2002) proposed deviation measure to quantify risk. Deviation measure is a function to evaluate financial risk in a different method than a general risk measure. A function $\mathcal{D} : \mathcal{L}^2 \to (0, +\infty)$ is a deviation risk measure if

- (D1): $\mathcal{D}(X + C) = \mathcal{D}(X)$ for all R and constants C, shift-invariant;
- (D2): $\mathcal{D}(C) = 0$ for constant C, but $\mathcal{D}(R) > 0$ for nonconstant R;
- (D3): $\mathcal{D}((1-\lambda)R + \lambda R) \leq (1-\lambda)\mathcal{D}(R) + \lambda\mathcal{D}(R)$ for $\lambda \in (0, 1)$ convexity;
- (D4): $\mathcal{D}(R + R') \leq \mathcal{D}(R) + \mathcal{D}(R')$ for all R and R';
- (D5): $\mathcal{D}(R) \leq d$ when $\| R^k - R \|_2 \to 0$ with $\mathcal{D}(R^k) \leq d$ (closedness).

There is a one-to-one correspondence between the deviation measures and risk measure as follow:

$$\mathcal{D}(R) = \mathcal{R}(R - E(R)) \tag{5.15}$$

$$\mathcal{R}(R) = \mathcal{D}(R) + E(R) \tag{5.16}$$

additionally, $\mathcal{R}$ is coherent $\Leftrightarrow$ $\mathcal{D}$ is coherent.

Standard deviation is a special case of deviation measure and it is symmetric. It is noteworthy to be aware that the deviation measure, which serves to evaluate the risk, is not the same thing as risk measure.

The c-VaR deviation measure and c-CVaR deviation measure, referring to Rockafellar et al. (2002), are defined as

$$VaR_c^{\triangle}(R) = VaR_c(R - E(R)) \tag{5.17}$$

and

$$CVaR_c^{\triangle}(R) = CVaR_c(R - E(R)) \tag{5.18}$$

as shown in Figure 5.3.

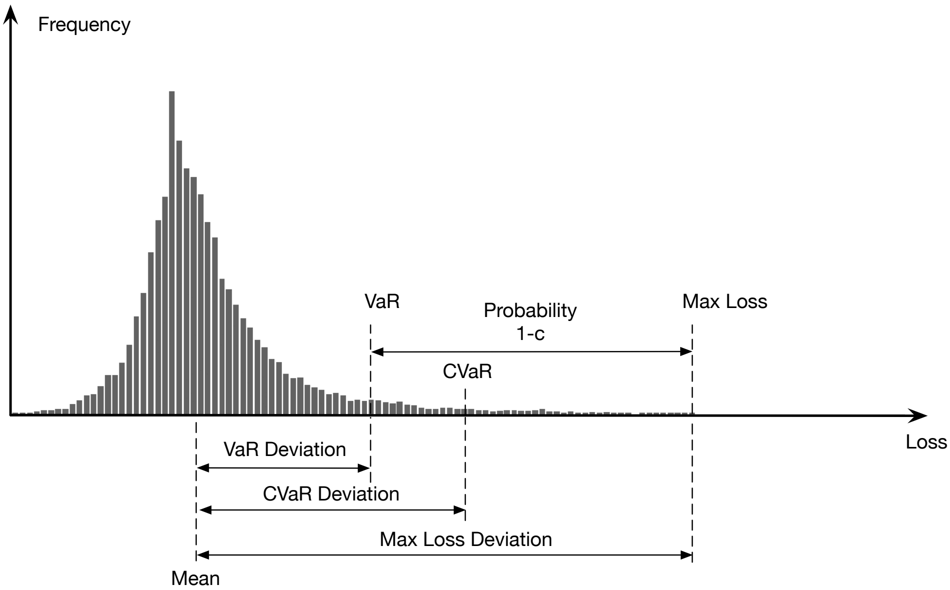

Figure 5.3: VaR deviation and CVaR deviation

The portfolio optimisation problem of CPT with a VaR deviation constraint can be written as

$$
\begin{aligned}
\max \quad & V(\boldsymbol{x}) && (5.19)\\
s.t. \quad & VaR_c^{\triangle} < l && (5.20)\\
& \sum_{i=1}^{n} x_i = 1, \\
& x_i \geq 0, i = 1, 2 \ldots, n
\end{aligned}
$$

where $V(\boldsymbol{x})$ is the equation (4.4).

Then the portfolio optimisation problem of CPT with a CVaR deviation constraint can be written as

$$
\begin{aligned}
\max \quad & V(\boldsymbol{x}) && (5.21)\\
s.t. \quad & CVaR_c^{\triangle} < l && (5.22)\\
& \sum_{i=1}^{n} x_i = 1, \\
& x_i \geq 0, i = 1, 2 \ldots, n
\end{aligned}
$$

where $V(\boldsymbol{x})$ is the equation (4.4).

5.3 Numerical experiments

This chapter tends to solve the problems (5.8), (5.9), (5.10),(5.11), and (5.12) using data from chapter 3. As for problems (5.19) and (5.21), we use the data presented in chapter 4.

The multivariate normal distribution is considered in this section, with the VaR constraints being given with 0.95 and 0.99 confidence level,i.e.$VaR_{0.95} < 0.01$ and $VaR_{0.99} < 0.01$. The following conclusions can be drawn according to the results shown in Table 5.1 by comparing to Table 3.4 provided in chapter 3. Our findings suggest that investment behaviors of CPT investors have changed significantly and CPT investors without constraint of VaR do not choose risk-free assets while CPT investors with constraint of VaR invest risk-free assets. Another important findings further reveal that CPT values were significantly reduced to CPT investor with VaR constraints. Compared with the cases without VaR constraints, the portfolio returns are reduced. It shows that risk reduction comes at the cost of returns.

Table 5.1: Results for solution of CPT function under VaR constraints

	γ	δ	CV	FR	x_1^*	x_2^*	x_3^*	x_4^*	x_5^*	c
α=0.36	0.56	0.56	0.1169	0.0299	0.2805	0	0.2265	0	0.4930	0.99
β=0.24	0.56	0.56	0.1170	0.0299	0.2809	0	0.2268	0	0.4923	0.95
	0.61	0.69	0.1412	0.0299	0.2825	0	0.2245	0	0.4930	0.99
	0.61	0.69	0.1412	0.0299	0.2789	0	0.2288	0	0.4923	0.95
α=0.50	0.56	0.56	0.0733	0.0299	0.2745	0	0.2328	0	0.4927	0.99
β=0.50	0.56	0.56	0.0733	0.0299	0.2734	0	0.2346	0	0.4920	0.95
	0.61	0.69	0.0810	0.0299	0.2804	0	0.2266	0	0.4930	0.99
	0.61	0.69	0.0810	0.0299	0.2847	0	0.2228	0	0.4925	0.95
α=0.88	0.56	0.56	0.0138	0.0299	0.2731	0	0.2343	0	0.4926	0.99
β=0.88	0.56	0.56	0.0138	0.0299	0.2770	0	0.2308	0	0.4922	0.95
	0.61	0.69	0.0149	0.0299	0.2770	0	0.2302	0	0.4928	0.99
	0.61	0.69	0.0149	0.0299	0.2881	0	0.2192	0	0.4927	0.95

Note: For each set of parameters, the results are the same over the 100 executions under population sizes of 30, 50, and 100.
CV stands for CPT Value; FR stands for Final Return; x_1^*, x_2^*, x_3^*, x_4^* and x_5^* stand for optimal investment ratios for stock Disney, Pfizer, Altria, Intel and riskless asset respectively.
$\lambda = 2.55$ and c represent the confidence level.

CVaR constraints are presented at given confidences with 0.95 and 0.99, assuming that the $CVaR_{0.95} < 0.005$ and $CVaR_{0.99} < 0.005$ respectively. By comparing the results of Table 5.2 and Table 3.3 provided in chapter 3, we found that investment behaviors of CPT investors have changed significantly. CPT investors without constraint of CVaR do not choose risk-free assets, while CPT investors with CVaR constraint will invest risk-free assets. And investment proportion on risk-free asset with constraint of CVaR is greater than investment proportion on risk-free asset with constraint of VaR, and that for the same CPT investor，the CPT value under CVaR constraint is much less than the CPT value without CVaR constraint.

From Table 5.1 and 5.2, we noticed that portfolio returns have decreased significantly once risk constraints are added, demonstrating that reduced portfolio returns are the cost of avoiding the potential risk.

Table 5.2: Results for solution of CPT function under CVaR constraints

	γ	δ	CV	FR	x_1^*	x_2^*	x_3^*	x_4^*	x_5^*	c
α=0.36	0.56	0.56	0.0883	0.0249	0.1387	0	0.1141	0	0.7472	0.99
β=0.24	0.56	0.56	0.0872	0.0248	0.1329	0	0.1118	0	0.7553	0.95
	0.61	0.69	0.1083	0.0249	0.1369	0	0.1160	0	0.7471	0.99
	0.61	0.69	0.1070	0.0248	0.1392	0	0.1053	0	0.7555	0.95
α=0.50	0.56	0.56	0.0517	0.0249	0.1366	0	0.1163	0	0.7471	0.99
β=0.50	0.56	0.56	0.0509	0.0248	0.1421	0	0.1023	0	0.7556	0.95
	0.61	0.69	0.0572	0.0249	0.1324	0	0.1208	0	0.7468	0.99
	0.61	0.69	0.0563	0.0248	0.1405	0	0.1040	0	0.7555	0.95
α=0.88	0.56	0.56	0.0075	0.0249	0.1326	0	0.1206	0	0.7468	0.99
β=0.88	0.56	0.56	0.0072	0.0248	0.1179	0	0.1275	0	0.7546	0.95
	0.61	0.69	0.0081	0.0249	0.1344	0	0.1187	0	0.7469	0.99
	0.61	0.69	0.0078	0.0248	0.1407	0	0.1037	0	0.7556	0.95

Note: For each set of parameters, the results are the same over the 100 executions under population sizes of 30, 50, and 100.
CV stands for CPT Value; FR stands for Final Return; x_1^*, x_2^*, x_3^*, x_4^* and x_5^* stand for optimal investment ratios for stock Disney, Pfizer, Altria, Intel and riskless asset respectively.
$\lambda = 2.55$ and c represent the confidence level.

A constrain of the proportion of riskless investment, $x_5^* \in (0.2, 0.4)$, is imposed on model (5.10). As shown in Table 5.3, the portfolio return is less than 0.20, showing the return of riskless asset. Consequently, based on the model (5.10), a constraint of portfolio return is added, i.e., $\mathring{r} > 0.2$ in model 5.11. The results also suggest that CPT investors reduce the proportion of riskless asset under the constraint of final portfolio

Table 5.3: Results for solution of CPT with constraints of riskless asset

Model	CV	FR	x_1^*	x_2^*	x_3^*	x_4^*	x_5^*	c
(5.10)	0.0279	0.0183	0.1510	0.2295	0.1265	0.1319	0.3612	\
(5.11)	0.0275	0.0225	0.2358	0.1646	0.1216	0.1742	0.3037	\
(5.12)	0.0146	0.0145	0.1896	0.3251	0.0877	0.0605	0.3372	0.95
(5.12)	0.0147	0.0123	0.0399	0.3483	0.1716	0.0555	0.3847	0.99

Note: For each set of parameters, the results are the same over the 100 executions under population sizes of 30, 50, and 100.
CV stands for CPT Value; FR stands for Final Return; x_1^*, x_2^*, x_3^*, x_4^* and x_5^* stand for optimal investment ratios for stock Disney, Pfizer, Altria, Intel and riskless asset respectively.
$\lambda = 2.55$ and c represent the confidence level.

return.

Table 5.4 also gives the results of model 5.12, providing the constrain with both riskless assets and VaR risk measure at 0.95 and 0.99 confidence level.

As can be seen in the Table 5.4, the proportion of investing in x_1 is greatly reduced under VaR constraint with the 0.99 confidence level, compared with the 0.95 confidence level.

Table 5.4: CPT value, portfolio return and optimal solution with VaR deviation measure at c=0.95 and l=0.004

	r_f	CPT VALUE	Return	DIS^*	GE^*	IBM^*
SB	0.003	0.0174	0.0137	0.6640	0.2934	0.0426
	0.004	0.0158	0.0137	0.6640	0.2926	0.0433
	0.005	0.0135	0.0131	0.6050	0.1814	0.2136
	0.006	0.0080	0.0123	0.4442	0.3102	0.2455
NBB	0.003	0.0180	0.0141	0.7104	0.2881	0.0015
	0.004	0.0164	0.0141	0.7104	0.2881	0.0015
	0.005	0.0146	0.0139	0.6755	0.3245	0
	0.006	0.0114	0.0128	0.4706	0.4428	0.0866
MBB	0.003	0.0215	0.0160	1	0	0
	0.004	0.0200	0.0160	1	0	0
	0.005	0.0185	0.0160	1	0	0
	0.006	0.0169	0.0160	1	0	0

The VaR and CVaR deviation constraints associate to confidence level c and let $VaR_{0.95}^{\triangle} < 0.004$,$VaR_{0.99}^{\triangle} < 0.004$, $CVaR_{0.95}^{\triangle} < 0.009$ and $CVaR_{0.99}^{\triangle} < 0.009$. It finds that there is a significantly affect the behavior of investors by introducing the risk con-

Table 5.5: CPT value, portfolio return and optimal solution with VaR deviation measure at c=0.99 and l=0.004

	r_f	CPT VALUE	Return	DIS^*	GE^*	IBM^*
SB	0.003	0.0146	0.0117	0.3327	0.3555	0.3118
	0.004	0.0130	0.0117	0.3330	0.3555	0.3115
	0.005	0.0114	0.0117	0.3330	0.3555	0.3115
	0.006	0.0056	0.0116	0.3048	0.4319	0.2633
NBB	0.003	0.0154	0.0123	0.4426	0.2487	0.3087
	0.004	0.0139	0.0123	0.4427	0.2489	0.3084
	0.005	0.0123	0.0123	0.4422	0.2491	0.3087
	0.006	0.0103	0.0120	0.3629	0.4137	0.2234
MBB	0.003	0.0215	0.0160	1	0	0
	0.004	0.0200	0.0160	1	0	0
	0.005	0.0185	0.0160	1	0	0
	0.006	0.0169	0.0160	1	0	0

Table 5.6: CPT value, portfolio return and optimal solution with CVaR deviation measure at c=0.95 and l=0.009

	r_f	CPT VALUE	Return	DIS^*	GE^*	IBM^*
SB	0.003	0.0194	0.0154	0.9999	0.0001	0
	0.004	0.0165	0.0143	0.8048	0.0912	0.1040
	0.005	0.0135	0.0131	0.6048	0.1816	0.2136
	0.006	0.0080	0.0122	0.4230	0.3422	0.2348
NBB	0.003	0.0191	0.0150	0.9024	0.0838	0.0138
	0.004	0.0163	0.0141	0.7451	0.1315	0.1234
	0.005	0.0140	0.0135	0.5895	0.4105	0
	0.006	0.0114	0.0128	0.4708	0.4418	0.0874
MBB	0.003	0.0135	0.0106	0.0077	0.9923	0
	0.004	0.0120	0.0106	0.0077	0.9923	0
	0.005	0.0103	0.0106	0.0077	0.9923	0
	0.006	0.0087	0.0106	0.0077	0.9923	0

Table 5.7: CPT value, portfolio return and optimal solution with CVaR deviation measure at c=0.99 and l=0.009

	r_f	CPT VALUE	Return	DIS^*	GE^*	IBM^*
SB	0.003	0.0194	0.0154	0.9999	0.0001	0
	0.004	0.0165	0.0143	0.8048	0.0912	0.1040
	0.005	0.0135	0.0131	0.6048	0.1816	0.2136
	0.006	0.0080	0.0122	0.4442	0.3102	0.2455
NBB	0.003	0.0191	0.0151	0.9075	0.0925	0
	0.004	0.0169	0.0145	0.7914	0.2086	0
	0.005	0.0146	0.0139	0.6755	0.3246	0
	0.006	0.0114	0.0128	0.4708	0.4418	0.0874
MBB	0.003	0.0141	0.0110	0.0763	0.9237	0
	0.004	0.0126	0.0110	0.0763	0.9237	0
	0.005	0.0109	0.0110	0.0763	0.9237	0
	0.006	0.0093	0.0110	0.0763	0.9237	0

straint in certain scenarios. As shown in Figure 5.4-5.7, the more stringent requirements for risk control, the more decentralized investment allocation, but the combination of investment income will be reduced accordingly, being consistent with the intuition, Thus, the VaR loses its effect as risk constraint under certain scene.

5.4 Summary

The chapter extends description of chapters 3 and 4 by using of VaR/CVaR risk measures and VaR/CVaR deviation measures for portfolio optimization under CPT. As far as is known, no paper would give the definitive results in the CPT optimization problem with constraint as mentioned in chapters 4 and 5. This chapter presents two kinds of model for optimizing portfolio under CPT with risk and deviation constraints using multivariate normal distribution and Bootstrap scenarios simulation.

This chapter discusses and compares the CPT objective function without risk or deviation constraints and with risk or deviation constraints. It found that CPT investor with constraints of risk and deviation significantly changed their investment behavior. Moreover, due to the constraints of risk and deviation, the CPT value decreased and the investment income declined.

Chapter 6

Summary

There has recently been increased research interest in CPT, but most studies have not taken portfolio optimization into account. It has been demonstrated that the optimization problem of CPT is a non-convex, non-concave, and non-smooth function. Consequently, it exceeds the capabilities of standard methods in optimization. In this dissertation, several single-period portfolio selection models were proposed for solving the portfolio choice problem under CPT in financial market. These study results provided several references for further research into the portfolio selection problem under CPT.

6.1 Contributions

To my knowledge, there has been little research on portfolio choice under CPT, save for the Pirvu and Schulze (2012) and Grishina et al. (2017), those studies have focused on experimental study or a market consisting of one risky asset and one riskless asset.

The main contributions in this dissertation are:

1. An adaptive real coded genetic algorithm was proposed to solve the CPT model due to complexity. The real coded genetic algorithm is better suited to large-dimensional search space and the adaptive properties can improve the efficiency of searching. This dissertation also presents the first study of the influence of various parameters for value function and weighting functions on the CPT value and portfolio selection.

Grishina et al. (2017) proposed a genetic algorithm to solve the optimization problem of CPT with value function and objective probability function, which ignores the important property of probability distortion. We painstakingly examined nonlinear value function and probability weighting functions in this dissertation.

2. We proposed a method by coupling bootstrap scenario generation techniques with genetic algorithm to solve the optimization problem under CPT. The method can extend research to any probability distributions of portfolio returns, such as skewness or fat tails. In addition, this was the first study to consider the impact of different reference points on investment behavior under CPT. Besides this dissertation demonstrated numerically that greater expectations lead to greater disappointment.

3. This dissertation solved the portfolio optimization under CPT with risk constraints, deviation constraints, and other constraints. As far as we know, no papers focus on portfolio optimization under CPT coupled with constraints of VaR and CVaR, save for the Pirvu and Schulze (2012). Even though Pirvu and Schulze (2012) imposed VaR and CVaR as constraints for solving the ill-posedness problem, which was non-existence in our model. Moreover, Pirvu and Schulze (2012) did not consider how the constraints of VaR and CVaR will affect behavior of CPT investor, which was gap to fulfill in this dissertation.

6.2 Future Work

Further investigation is needed to compare the portfolio choices under EUT with those under CPT, and a discussion is expected to be provided about the implications of these differences on the decision-making behavior of perfect and bounded rationality. Because CPT can describe the behavior of bounded rational decision makers in a psychologically more realistic way, over the past decade, researchers in the field of behavioral economics have repeatedly considered how prospect theory is to be applied in economic settings; these efforts are now bearing fruit.

To the best of our knowledge, however, very few papers have studied the subject of comparing EUT and CPT. Sebora and Cornwall (1995) studied the implications for strategic decision makers under EUT and CPT. Bleichrodt et al. (2001) proposed a quantitative modification of standard utility elicitation procedures using the idea of CPT to correct for commonly observed violations of expected utility. Harrison and Rutström (2009) compared EUT with CPT on one wedding and a decent funeral.

Although the future work is to compare the results of portfolio optimization under CPT and EUT, it is difficult to say which is better than the other. EUT and CPT are two different theories based on different assumptions. EUT is originally developed as a normative theory about how rational decision makers can maximize their utility. CPT is conceived as a descriptive theory of how boundedly rational decision makers (real decision makers) ultimately achieve the most satisfying results. As Tversky and Kahneman (1986) stated, “The normative and the descriptive analyses of choice should be viewed as separate enterprises.”

Appendix A

Proofs

Proof of Proposition 1: According to the definitions of the value functions and weighing functions, they all are differentiable. We can obtain

$$\begin{aligned}&\int_0^{+\infty} v^+(t)dw^+(1-F_D(t))\\&= v^+(t)w^+(1-F_D(t))|_0^{+\infty}\\&-\int_0^{+\infty} w^+(1-F_D(t))dv^+(t)\end{aligned} \tag{A.1}$$

using integration by parts.

He and Zhou (2011) showed that, if the return on a portfolio follows a normal distribution and $|t|$ and $0 < F_D(t) < 1$ are sufficiently large, then (3.5) and (3.6) satisfy:

$$\begin{aligned}w^{+'}(F_D(t))f_D(t) &= O(|t|^{-2-\varepsilon})\\w^{-'}(F_D(t))f_D(t) &= O(|t|^{-2-\varepsilon})\end{aligned} \tag{A.2}$$

$$\begin{aligned}w^{+'}(1-F_D(t))f_D(t) &= O(|t|^{-2-\varepsilon})\\w^{-'}(1-F_D(t))f_D(t) &= O(|t|^{-2-\varepsilon})\end{aligned} \tag{A.3}$$

where $\varepsilon > 0$.

Furthermore, from L'Hopital's rule, we know that

$$\begin{aligned}&\lim_{t\to\infty} v^+(t)w^+(1-F_D(t))\\&= \lim_{t\to\infty} \frac{w^+(1-F_D(t))}{\frac{1}{v^+(t)}} = 0\end{aligned} \tag{A.4}$$

We can obtain that

$$\int_0^{+\infty} v^+(t)dw^+(1-F_D(t))$$
$$= -\int_0^{+\infty} w^+(1-F_D(t))dv^+(t) \tag{A.5}$$

Similarly, we have

$$\int_{-\infty}^{0} v^-(t)dw^+(F_D(t))$$
$$= -\int_{-\infty}^{0} w^-(F_D(t))dv^-(t) \tag{A.6}$$

Proof of Proposition 2: Hogg and Craig (1995) have shown that a linear transformation of multivariate normal random vectors has a multivariate normal distribution, namely, suppose that $\boldsymbol{R}$ has the distribution $N_n(\boldsymbol{\mu}, \Sigma)$ and $D = \boldsymbol{AR} + b$, where $\boldsymbol{A}$ is an $m \times n$ matrix and $b \in \boldsymbol{R}^m$. Then, D has the distribution $N_m(\boldsymbol{A\mu}, \boldsymbol{A}\Sigma\boldsymbol{A'})$. Proposition 1 can be proved when $m = 1$.

Poof of Proposition 3:

Assume that portfolio return R is a norm with mean μ_p and variance σ_p^2, then:

$$f(r) = \frac{1}{\sqrt{2\pi}\sigma_p} e^{-\frac{(r-\mu_p)^2}{2\sigma_p^2}} \tag{A.7}$$

and

$$z = \frac{r-\mu_p}{\sigma_p} \Leftrightarrow r = \sigma_p z + \mu_p \tag{A.8}$$

z is standard normal random variable and has density function as follow:

$$\phi(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} \tag{A.9}$$

CPT functions can be represented as follows:

$$V(R) = V(R^+) + V(R^-) \tag{A.10}$$
$$V(R^+) = -\int_0^{+\infty} v^+(r)dw^+(1-F_R(r)) \tag{A.11}$$
$$V(R^-) = \int_{-\infty}^{0} v^-(r)dw^-(F_R(r)) \tag{A.12}$$

and

$$
\begin{aligned}
V^+(R) &= \int_0^{+\infty} v^+(r) f_R(r) dr && \text{(A.13)}\\
&= \int_{-\frac{\mu_p}{\sigma_p}}^{+\infty} v^+(\sigma_p z + \mu) f_R(\sigma_p z + \mu_p) d(\sigma_p z + \mu_p) && \text{(A.14)}\\
&= \int_{-\frac{\mu_p}{\sigma_p}}^{+\infty} v^+(\sigma_p z + \mu_p) \frac{1}{\sqrt{2\pi}\sigma_p} e^{-\frac{(\sigma_p z + \mu - \mu_p)^2}{2\sigma_p^2}} \sigma_p dz && \text{(A.15)}\\
&= \int_{-\frac{\mu_p}{\sigma_p}}^{+\infty} v^+(\sigma_p z + \mu_p) \frac{1}{\sqrt{2\pi}\sigma_p} e^{-\frac{(\sigma_p z)^2}{2\sigma_p^2}} \sigma_p dz && \text{(A.16)}\\
&= \int_{-\frac{\mu_p}{\sigma_p}}^{+\infty} v^+(\sigma_p z + \mu_p) \frac{1}{\sqrt{2\pi}\sigma_p} e^{-\frac{z^2}{2}} \sigma_p dz && \text{(A.17)}\\
&= \int_{-\frac{\mu_p}{\sigma_p}}^{+\infty} v^+(\sigma_p z + \mu_p) \phi(z) dz && \text{(A.18)}
\end{aligned}
$$

It takes the derivative of $V^+(R)$ with respect to μ_p

$$
\frac{\partial V^+(R)}{\partial \mu_p} = \int_{-\frac{\mu_p}{\sigma_p}}^{+\infty} v^{+'}(\sigma_p z + \mu_p) \phi(z) dz > 0 \tag{A.19}
$$

Similarly,

$$
V^-(R) = \int_{-\infty}^{-\frac{\mu_p}{\sigma_p}} v^-(\sigma_p z + \mu_p) \phi(z) dz \tag{A.20}
$$

It takes the derivative of $V^-(R)$ with respect to μ_p

$$
\frac{\partial V^-(R)}{\partial \mu_p} = \int_{-\infty}^{-\frac{\mu_p}{\sigma_p}} v^{-'}(\sigma_p z + \mu_p) \phi(z) dz > 0 \tag{A.21}
$$

It takes the derivative of $V^+(R)$ with respect to σ_p

$$
\frac{\partial V^+(R)}{\partial \sigma_p} = \int_{-\frac{\mu_p}{\sigma_p}}^{+\infty} v^{+'}(\sigma_p z + \mu_p) \phi(z) z dz \tag{A.22}
$$

The derivative of the standard normal distribution is

$$
\phi'(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} \cdot (-z) = -z\phi(z) \tag{A.23}
$$

$$\frac{\partial V^{+}(R)}{\partial \sigma_p} = \int_{-\frac{\mu_p}{\sigma_p}}^{+\infty} v^{+'}(\sigma_p z + \mu_p)\{-\phi(z)\}' dz \tag{A.24}$$

$$= -[v^{+'}(\sigma_p z + \mu_p)\phi(z)]|_{z=-\frac{\mu_p}{\sigma_p}}^{z=+\infty} + \int_{-\frac{\mu_p}{\sigma_p}}^{+\infty} v^{+''}(\sigma_p z + \mu_p)\sigma_p \phi(z) dz \tag{A.25}$$

$$= v^{+'}(0)\phi(-\frac{\mu_p}{\sigma_p}) + \sigma_p \int_{-\frac{\mu_p}{\sigma_p}}^{+\infty} v^{+''}(\sigma_p z + \mu_p)\phi(z) dz \tag{A.26}$$

It is obvious that $v^{+'}(0)\phi(-\frac{\mu_p}{\sigma_p}) > 0$ and $\sigma_p \int_{-\frac{\mu_p}{\sigma_p}}^{+\infty} v^{+''}(\sigma_p z + \mu_p)\phi(z) < 0$
Similarly,

$$\frac{\partial V^{-}(R)}{\partial \sigma_p} = v^{-'}(0)\phi(-\frac{\mu_p}{\sigma_p}) + \sigma_p \int_{-\infty}^{-\frac{\mu_p}{\sigma_p}} v^{-''}(\sigma_p z + \mu_p)\phi(z) dz > 0 \tag{A.27}$$

Appendix B

Portfolio choice based on rationality

The efficient frontier can be obtained according to 3.22 and 3.23, as shown in Figure B.1.

Suppose the risk aversion parameter are 2, 3, and 4 in problem (2.13), which are represent three kinds of investors, the portfolio returns are 0.0395552，0.0395479，and 0.0395445. It confirms that risk-averse will lead to a lower expected return. The problem (2.14) shows that the optimal portfolio is the position being tangent to the efficient frontier.

The problem (2.15) is the case where riskless asset is introduced. According to mutual fund separation theorem, mean-variance efficient portfolios can be formed simply as a combination of holdings of the risk-free asset and holdings of a particular efficient fund that contains only risky assets. We give the result of tangency portfolio that is the market portfolio including all risky assets. The results are shown in the Table B.1.

Table B.1: Results for rational solution

Problem	μ_p	σ_p	x_1^*	x_2^*	x_3^*	x_4^*
2.13-2	0.0395552	0.0084773	0.5552	0	0.4448	0
2.13-3	0.0395479	0.0084743	0.5479	0	0.4521	0
2.13-4	0.0395445	0.0084733	0.5445	0	0.4555	0
2.14	0.0395429	0.0084729	0.5429	0	0.4571	0
2.15	0.0395524	0.0084760	0.5524	0	0.4476	0

Note: 2.13-2, 2.13-3 and 2.13-4 means that b is 2,3 and 4 in equation (2.13) respectively. x_1^*, x_2^*, x_3^*, and x_4^* stand for optimal investment ratios for stock Disney, Pfizer, Altria, and Intel respectively.

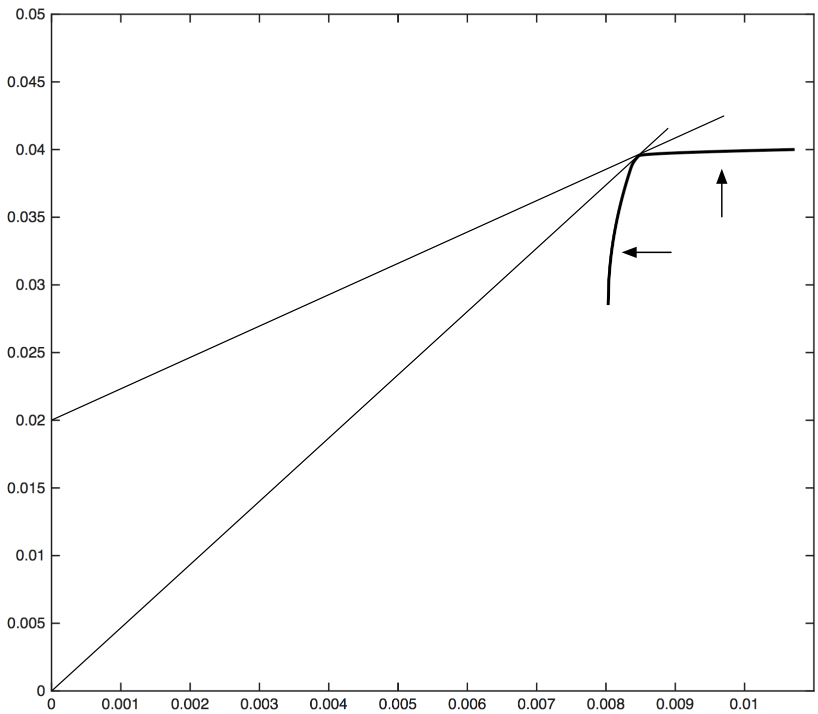

Figure B.1: The efficient frontier in chapter 3

Bibliography

Alexander, G. J. and Baptista, A. M. (2002). Economic implications of using a mean-var model for portfolio selection: A comparison with mean-variance analysis. *Journal of Economic Dynamics and Control*, 26(7):1159–1193.

Alexander, G. J. and Baptista, A. M. (2004). A comparison of var and cvar constraints on portfolio selection with the mean-variance model. *Management Science*, 50(9):1261–1273.

Ali, H. O. and Jilani, F. (2014). Mean-var model with stochastic volatility. *Procedia-Social and Behavioral Sciences*, 109:558–566.

Allais, M. (1953). Le comportement de l'homme rationnel devant le risque: critique des postulats et axiomes de l'école américaine. *Econometrica: Journal of the Econometric Society*, pages 503–546.

Artzner, P., Delbaen, F., Eber, J.-M., and Heath, D. (1999). Coherent measures of risk. *Mathematical finance*, 9(3):203–228.

Back, T. (1994). Selective pressure in evolutionary algorithms: A characterization of selection mechanisms. In *Evolutionary Computation, 1994. IEEE World Congress on Computational Intelligence., Proceedings of the First IEEE Conference on*, pages 57–62. IEEE.

Baixauli-Soler, J. S., Alfaro-Cid, E., and Fernandez-Blanco, M. O. (2011). Mean-var portfolio selection under real constraints. *Computational Economics*, 37(2):113–131.

Baker, J. E. (1985). Adaptive selection methods for genetic algorithms. In *Proceedings of an International Conference on Genetic Algorithms and their applications*, pages 101–111. Hillsdale, New Jersey.

Barberis, N. and Huang, M. (2008). Stocks as lotteries: The implications of probability weighting for security prices. *The American Economic Review*, 98(5):2066–2100.

Barberis, N., Huang, M., and Santos, T. (2001). Prospect theory and asset prices. *Quarterly Journal of Economics*, 116(1):1–53.

Barberis, N., Huang, M., and Thaler, R. H. (2006). Individual preferences, monetary gambles, and stock market participation: A case for narrow framing. *The American economic review*, 96(4):1069–1090.

Barberis, N. C. (2013). Thirty years of prospect theory in economics: A review and assessment. *The Journal of Economic Perspectives*, 27(1):173–195.

Basak, S. and Shapiro, A. (2001). Value-at-risk-based risk management: optimal policies and asset prices. *Review of Financial studies*, 14(2):371–405.

Baskar, S., Subbaraj, P., and Chidambaram, P. (2001). Application of genetic algorithms to unit commitment problem. *Journal of Institute of Engineering (India)*, 81:195–199.

Baskar, S., Subbaraj, P., and Rao, M. (2004). Performance of real-coded genetic algorithms with new genetic operators. *J. Inst. Eng.(India)*, 85:124–128.

Baskar, S., Subbaraj, P., Rao, M., and Tamilselvi, S. (2003). Genetic algorithms solution to generator maintenance scheduling with modified genetic operators. *IEE Proceedings-Generation, Transmission and Distribution*, 150(1):56–60.

Benartzi, S. and Thaler, R. H. (1995). Myopic loss aversion and the equity premium puzzle. *The Quarterly Journal of Economics*, 110:73–92.

Bernard, C. and Ghossoub, M. (2010). Static portfolio choice under cumulative prospect theory. *Mathematics and financial economics*, 2(4):277–306.

Bernoulli, D. (1954). Exposition of a new theory on the measurement of risk. *Econometrica: Journal of the Econometric Society*, pages 23–36.

Birnbaum, M. H. and Chavez, A. (1997). Tests of theories of decision making: Violations of branch independence and distribution independence. *Organizational Behavior and human decision Processes*, 71(2):161–194.

Bleichrodt, H. and Pinto, J. L. (2000). A parameter-free elicitation of the probability weighting function in medical decision analysis. *Management science*, 46(11):1485–1496.

Bleichrodt, H., Pinto, J. L., and Wakker, P. P. (2001). Making descriptive use of prospect theory to improve the prescriptive use of expected utility. *Management science*, 47(11):1498–1514.

Blondel, S. (2002). Testing theories of choice under risk: Estimation of individual functionals. *Journal of Risk and Uncertainty*, 24(3):251–265.

Blum, S., Puisa, R., Riedel, J., and Wintermantel, M. (2001). Adaptive mutation strategies for evolutionary algorithms. In *The Annual Conference: EVEN at Weimarer Optimierungsund Stochastiktage*.

Boudt, K., Carl, P., and Peterson, B. G. (2013). Asset allocation with conditional value-at-risk budgets. *Ssrn Electronic Journal*, 15(3):39–68.

Bradley, S. P. and Crane, D. B. (1972). A dynamic model for bond portfolio management. *Management Science*, 19(2):139–151.

Bühlmann, P. and Künsch, H. R. (1995). The blockwise bootstrap for general parameters of a stationary time series. *Scandinavian Journal of Statistics*, pages 35–54.

Camerer, C. F. and Ho, T.-H. (1994). Violations of the betweenness axiom and nonlinearity in probability. *Journal of risk and uncertainty*, 8(2):167–196.

Campbell, R., Huisman, R., and Koedijk, K. (2001). Optimal portfolio selection in a value-at-risk framework. *Journal of Banking & Finance*, 25(9):1789–1804.

Carlstein, E. (1986). The use of subseries values for estimating the variance of a general statistic from a stationary sequence. *The Annals of Statistics*, pages 1171–1179.

Chance, D. M. (2004). *Introduction to derivatives and risk management*. Harcourt College Publishers,.

Chang, T.-J., Meade, N., Beasley, J. E., and Sharaiha, Y. M. (2000). Heuristics for cardinality constrained portfolio optimisation. *Computers & Operations Research*, 27(13):1271–1302.

Chelouah, R. and Siarry, P. (2003). Genetic and nelder–mead algorithms hybridized for a more accurate global optimization of continuous multiminima functions. *European Journal of Operational Research*, 148(2):335–348.

Coelho, L. A. G. et al. (2014). Portfolio selection optimization under cumulative prospect theory–a parameter sensibility analysis. *CEFAGE-UE working Papers*.

Consigli, G. (2002). Tail estimation and mean–var portfolio selection in markets subject to financial instability. *Journal of Banking & Finance*, 26(7):1355–1382.

Crow, J. F. and Kimura, M. (1979). Efficiency of truncation selection. *Proceedings of the National Academy of Sciences*, 76(1):396–399.

De Giorgi, E., Hens, T., and Mayer, J. (2007). Computational aspects of prospect theory with asset pricing applications. *Computational Economics*, 29(3):267–281.

Deb, K. and Beyer, H.-G. (1999). Self-adaptation in real-parameter genetic algorithms with simulated binary crossover. In *Proceedings of the 1st Annual Conference on Genetic and Evolutionary Computation-Volume 1*, pages 172–179. Morgan Kaufmann Publishers Inc.

Deb, K. and Goyal, M. (1996). A combined genetic adaptive search (geneas) for engineering design. *Computer Science and informatics*, 26:30–45.

Deep, K. and Thakur, M. (2007). A new crossover operator for real coded genetic algorithms. *Applied mathematics and computation*, 188(1):895–911.

Dowd, K. (2003). *An introduction to market risk measurement*. John Wiley & Sons.

Dowd, K. K. (2002). *An introduction to market risk measurement /*. J. Wiley,.

Duffie, D. (2010). *Dynamic asset pricing theory*. Princeton University Press.

Efron, B. (1992). Bootstrap methods: another look at the jackknife. In *Breakthroughs in Statistics*, pages 569–593. Springer.

Efron, B. and Tibshirani, R. J. (1993). *An Introduction to the Bootstrap*. New York Chapman Hall.

Ellsberg, D. (1961). Risk, ambiguity, and the savage axioms. *The quarterly journal of economics*, pages 643–669.

Estrada, J. (2007). Mean-semivariance behavior: Downside risk and capital asset pricing. *International Review of Economics & Finance*, 16(2):169–185.

Fennema, H. and Van Assen, M. (1998). Measuring the utility of losses by means of the tradeoff method. *Journal of risk and uncertainty*, 17(3):277–296.

Föllmer, H. and Schied, A. (2011). *Stochastic finance: an introduction in discrete time*. Walter de Gruyter.

Franke, J., Neumann, M. H., and Stockis, J.-P. (2004). Bootstrapping nonparametric estimators of the volatility function. *Journal of Econometrics*, 118(1):189–218.

Gaivoronski, A. A. and Pflug, G. (1999). Finding optimal portfolios with constraints on value at risk. In *Proceedings III Stockholm seminar on risk behavior and risk management*.

Glimcher, P. W. and Fehr, E. (2013). *Neuroeconomics: Decision making and the brain*. Academic Press.

Goldberg, D. E. (1989). Genetic algorithms in search, optimization and machine learning. xiii(7):2104–2116.

Goldberg, D. E. and Deb, K. (1991). A comparative analysis of selection schemes used in genetic algorithms. *Foundations of genetic algorithms*, 1:69–93.

Gonzalez, R. and Wu, G. (1999). On the shape of the probability weighting function. *Cognitive psychology*, 38(1):129–166.

Graham, R. L., Knuth, D. E., and Patashnik, O. (1994). Concrete math. a foundation for computer science.

Grishina, N., Lucas, C., and Date, P. (2017). Prospect theory–based portfolio optimization: an empirical study and analysis using intelligent algorithms. *Quantitative Finance*, 17(3):353–367.

Hall, P. (1985). Resampling a coverage pattern. *Stochastic processes and their applications*, 20(2):231–246.

Hancock, P. (1994). An empirical comparison of selection methods in evolutionary algorithms. *Evolutionary Computing*, pages 80–94.

Härdle, W., Müller, M., Sperlich, S., and Werwatz, A. (2012). *Nonparametric and semiparametric models*. Springer Science & Business Media.

Harrison, G. W. and Rutström, E. E. (2009). Expected utility theory and prospect theory: One wedding and a decent funeral. *Experimental economics*, 12(2):133–158.

He, X. D. and Zhou, X. Y. (2011). Portfolio choice under cumulative prospect theory: An analytical treatment. *Management Science*, 57(2):315–331.

Herrera, F., Lozano, M., and Verdegay, J. L. (1998). Tackling real-coded genetic algorithms: Operators and tools for behavioural analysis. *Artificial intelligence review*, 12(4):265–319.

Hoffmeister, F. and Bäck, T. (1991). *Genetic Algorithms and evolution strategies: Similarities and differences*. Springer Berlin Heidelberg.

Hogan, W. W. and Warren, J. M. (1974). Toward the development of an equilibrium capital-market model based on semivariance. *Journal of Financial and Quantitative Analysis*, pages 1–11.

Hogg, R. V. and Craig, A. T. (1995). *Introduction to mathematical statistics*. Upper Saddle River, New Jersey: Prentice Hall.

Holland, J. H. (1975). Adaptation in natural and artificial systems. *Control and Artificial Intelligence University of Michigan Press*, 6(2):126–137.

Horcher, K. A. (2011). *Essentials of financial risk management*, volume 32. John Wiley & Sons.

Hrstka, O. and Kučerová, A. (2004). Improvements of real coded genetic algorithms based on differential operators preventing premature convergence. *Advances in Engineering Software*, 35(3):237–246.

Hu, W. and Kercheval, A. N. (2010). Portfolio optimization for student t and skewed t returns. *Quantitative Finance*, 10(1):91–105.

Huang, C. and Litzenberger, R. H. (1988). Foundations of financial economics. *North-Holland Publishing Co., New York*.

Huimei, Y. (2000). Adaptive crossover rate and mutation rate in genetic algorithms. *Journal of Capital Normal University (Natural Science Edition)*, 21(3):14–20.

Hull, J. C. (2008). *Options, futures, and other derivat&es*. Pearson Prentice Hall, Upper Saddle River, NJ.

Ingersoll, J. (2008). Non-monotonicity of the tversky-kahneman probability-weighting function: A cautionary note. *European Financial Management*, 14(3):385–390.

Ingersoll, J. E. (1987). *Theory of financial decision making*, volume 3. Rowman & Littlefield.

Irving, F. (1906). The nature of capital and income. *Macmillam*, pages 1–21.

Jia, J. and Dyer, J. S. (1996). A standard measure of risk and risk-value models. *Management Science*, 42(12):1691–1705.

Jin, H., Markowitz, H., and Yu Zhou, X. (2006). A note on semivariance. *Mathematical Finance*, 16(1):53–61.

Jin, H. and Yu Zhou, X. (2008). Behavioral portfolio selection in continuous time. *Mathematical Finance*, 18(3):385–426.

Jorion, P. (2000). *Value at risk : the new benchmark for managing financial risk / 2nd. ed*. McGraw-Hill.

Jorion, P. (2006). Value at risk, 3rd ed.: The new benchmark for managing financial risk. *McGraw-Hill Education*.

JPMorgan (1994). Introduction to riskmetrics. *New York: JP Morgan*.

JPMorgan (1997). Creditmetrics-technical document. *JP Morgan, New York*.

Kaelo, P. and Ali, M. (2007). Integrated crossover rules in real coded genetic algorithms. *European Journal of Operational Research*, 176(1):60–76.

Kahneman, D. and Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica: Journal of the econometric society*, pages 263–291.

Karatzas, I., Shreve, S. E., Karatzas, I., and Shreve, S. E. (1998). *Methods of mathematical finance*, volume 1. Springer.

Karni, E. (2014). Axiomatic foundations of expected utility and subjective probability. *Handbook of the Economics of Risk and Uncertainty*, 1.

Kirkpatrick, S., Gelatt, C. D., Vecchi, M. P., et al. (1983). Optimization by simulated annealing. *science*, 220(4598):671–680.

Kleindorfer, P. R. and Li, L. (2005). Multi-period var-constrained portfolio optimization with applications to the electric power sector. *The Energy Journal*, pages 1–26.

Kouwenberg, R. and Zenios, S. A. (2006). Stochastic programming models for asset liability management. *Handbook of asset and liability management*, 1:253–303.

Kreiss, J.-P. and Lahiri, S. (2012). Bootstrap methods for time series. *Handbook of Statistics: Time Series Analysis: Methods and Applications*, 30(1).

Krokhmal, P., Palmquist, J., and Uryasev, S. (2002). Portfolio optimization with conditional value-at-risk objective and constraints. *Journal of risk*, 4:43–68.

Kroll, Y., Levy, H., and Markowitz, H. M. (1984). Mean-variance versus direct utility maximization. *The Journal of Finance*, 39(1):47–61.

Kunsch, H. R. (1989). The jackknife and the bootstrap for general stationary observations. *The annals of Statistics*, pages 1217–1241.

Lahiri, S. N. (1999). Theoretical comparisons of block bootstrap methods. *Annals of Statistics*, pages 386–404.

Levy, H. (2011). *The capital asset pricing model in the 21st century: analytical, empirical, and behavioral perspectives*. Cambridge University Press.

Levy, H. and Levy, M. (2004). Prospect theory and mean-variance analysis. *Review of Financial Studies*, 17(4):1015–1041.

Levy, H. and Markowitz, H. M. (1979). Approximating expected utility by a function of mean and variance. *The American Economic Review*, pages 308–317.

Lin, W.-Y., Lee, W.-Y., and Hong, T.-P. (2003). Adapting crossover and mutation rates in genetic algorithms. *J. Inf. Sci. Eng.*, 19(5):889–903.

Linsmeier, T. J. and Pearson, N. D. (2000). Value at risk. *Financial Analysts Journal*, pages 47–67.

Liu, R. Y. and Singh, K. (1992). Moving blocks jackknife and bootstrap capture weak dependence. *Exploring the limits of bootstrap*, 225:248.

Lopes, L. L. (1987). Between hope and fear: The psychology of risk. *Advances in experimental social psychology*, 20:255–295.

Luce, R. D. (1991). Rank-and sign-dependent linear utility models for binary gambles. *Journal of Economic Theory*, 53(1):75–100.

Luce, R. D. (2001). Reduction invariance and prelec's weighting functions. *Journal of Mathematical Psychology*, 45(1):167–179.

Markowitz, H. (1952). Portfolio selection*. *The journal of finance*, 7(1):77–91.

Markowitz, H. (1959). Portfolio selection: efficient diversification of investments. *Cowies Foundation Monograph*, (16).

Markowitz, H. (2014). Mean–variance approximations to expected utility. *European Journal of Operational Research*, 234(2):346–355.

Marschak, J. (1938). Money and the theory of assets. *Econometrica, Journal of the Econometric Society*, pages 311–325.

Mausser, H. and Rosen, D. (1999). Beyond var: from measuring risk to managing risk. In *Computational Intelligence for Financial Engineering*, pages 163–178.

Merton, R. C. (1969). Lifetime portfolio selection under uncertainty: The continuous-time case. *The review of Economics and Statistics*, pages 247–257.

Merton, R. C. and Samuelson, P. A. (1990). Continuous-time finance.

Meyer, J. (1987). Two-moment decision models and expected utility maximization. *The American Economic Review*, pages 421–430.

Michalewicz, Z. (1994). Genetic algorithms+ data structures= evolution programs. *Computational Statistics and Data Analysis*, 24(3):372–373.

Nawrocki, D. N. (1999). A brief history of downside risk measures. *The Journal of Investing*, 8(3):9–25.

Neumann, J. v., Morgenstern, O., et al. (1944). *Theory of games and economic behavior*, volume 60. Princeton university press Princeton.

Pennacchi, G. (2007). *Theory of asset pricing*. Pearson.

Pflug, G. C. (2000). Some remarks on the value-at-risk and the conditional value-at-risk. In *Probabilistic constrained optimization*, pages 272–281. Springer.

Pirvu, T. A. and Schulze, K. (2012). Multi-stock portfolio optimization under prospect theory. *Mathematics and financial economics*, 6(4):337–362.

Porter, R. B. (1974). Semivariance and stochastic dominance: A comparison. *The American Economic Review*, 64(1):200–204.

Quiggin, J. (1982). A theory of anticipated utility. *Journal of Economic Behavior & Organization*, 3(4):323–343.

Ranković, V., Drenovak, M., Stojanović, B., Kalinić, Z., and Arsovski, Z. (2014). The mean-value at risk static portfolio optimization using genetic algorithm. *Computer Science and Information Systems*, 11(1):89–109.

Rockafellar, R. T. (2007). Coherent approaches to risk in optimization under uncertainty. In *OR Tools and Applications: Glimpses of Future Technologies*, pages 38–61. Informs.

Rockafellar, R. T. and Uryasev, S. (2000). Optimization of conditional value-at-risk. *Journal of risk*, 2:21–42.

Rockafellar, R. T. and Uryasev, S. (2002). Conditional value-at-risk for general loss distributions. *Journal of banking & finance*, 26(7):1443–1471.

Rockafellar, R. T., Uryasev, S. P., and Zabarankin, M. (2002). Deviation measures in risk analysis and optimization. *Social Science Electronic Publishing*.

Ross, S. M. (2011). *An elementary introduction to mathematical finance*. Cambridge University Press.

Roy, A. D. (1952). Safety first and the holding of assets. *Econometrica: Journal of the Econometric Society*, pages 431–449.

Rubinstein, M. (2002). Markowitz's "portfolio selection": A fifty-year retrospective. *The Journal of finance*, 57(3):1041–1045.

Rudolph, G. (1994). Convergence analysis of canonical genetic algorithms. *IEEE transactions on neural networks*, 5(1):96–101.

Samuelson, P. A. (1969). Lifetime portfolio selection by dynamic stochastic programming. *The review of economics and statistics*, pages 239–246.

Samuelson, W. and Zeckhauser, R. (1988). Status quo bias in decision making. *Journal of risk and uncertainty*, 1(1):7–59.

Sarykalin, S., Serraino, G., and Uryasev, S. (2008). Value-at-risk vs. conditional value-at-risk in risk management and optimization. *Tutorials in Operations Research. INFORMS, Hanover, MD*, pages 270–294.

Schmidt, U. and Zank, H. (2008). Risk aversion in cumulative prospect theory. *Management Science*, 54(1):208–216.

Schoemaker, P. J. (1982). The expected utility model: Its variants, purposes, evidence and limitations. *Journal of economic literature*, pages 529–563.

Sebora, T. C. and Cornwall, J. R. (1995). Expected utility theory vs. prospect theory: Implications for strategic decision makers. *Journal of Managerial Issues*, pages 41–61.

Sharpe, W. F. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk*. *The journal of finance*, 19(3):425–442.

Sharpe, W. F. (1966). Mutual fund performance. *The Journal of business*, 39(1):119–138.

Sheng, Z., Benshan, S., and Zhongping, W. (2012). Analysis of mean-var model for financial risk control. *Systems Engineering Procedia*, 4:40–45.

Singh, K. (1981). On the asymptotic accuracy of efron's bootstrap. *The Annals of Statistics*, pages 1187–1195.

Spall, J. C. (2005). *Introduction to stochastic search and optimization: estimation, simulation, and control*, volume 65. John Wiley & Sons.

Spanier, J. and Oldham, K. B. (1987). *An atlas of functions*. Taylor & Francis/Hemisphere.

Srinivas, M. and Patnaik, L. M. (1994). Adaptive probabilities of crossover and mutation in genetic algorithms. *IEEE Transactions on Systems, Man, and Cybernetics*, 24(4):656–667.

Stracca, L. (2002). The optimal allocation of risks under prospect theory. *Social Science Electronic Publishing*, 27(2):697–711.

Subbaraj, P., Rengaraj, R., and Salivahanan, S. (2009). Enhancement of combined heat and power economic dispatch using self adaptive real-coded genetic algorithm. *Applied Energy*, 86(6):915–921.

Subbaraj, P., Rengaraj, R., and Salivahanan, S. (2011). Enhancement of self-adaptive real-coded genetic algorithm using taguchi method for economic dispatch problem. *Applied Soft Computing*, 11(1):83–92.

Thaler, R. (1980). Toward a positive theory of consumer choice. *Journal of Economic Behavior & Organization*, 1(1):39–60.

Thierens, D. and Goldberg, D. (1994a). Convergence models of genetic algorithm selection schemes. *Parallel problem solving from nature—PPSN III*, pages 119–129.

Thierens, D. and Goldberg, D. (1994b). Elitist recombination: An integrated selection recombination ga. In *Evolutionary Computation, 1994. IEEE World Congress on*

Computational Intelligence., Proceedings of the First IEEE Conference on, pages 508–512. IEEE.

Tian, R., Cox, S. H., Lin, Y., and Zuluaga, L. F. (2010). Portfolio risk management with cvar-like constraints. *North American Actuarial Journal*, 14(1):86–106.

Topaloglou, N., Vladimirou, H., and Zenios, S. A. (2002). Cvar models with selective hedging for international asset allocation. *Journal of Banking & Finance*, 26(7):1535–1561.

Tsao, C.-Y. (2010). Portfolio selection based on the mean–var efficient frontier. *Quantitative Finance*, 10(8):931–945.

Tsutsui, S. and Goldberg, D. E. (2001). Search space boundary extension method in real-coded genetic algorithms. *Information Sciences*, 133(3):229–247.

Tversky, A. and Kahneman, D. (1986). Rational choice and the framing of decisions. *Journal of business*, 59:251–278.

Tversky, A. and Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and uncertainty*, 5(4):297–323.

Whitley, L. D. et al. (1989). The genitor algorithm and selection pressure: Why rank-based allocation of reproductive trials is best. In *ICGA*, volume 89, pages 116–123.

Wu, G. and Gonzalez, R. (1996). Curvature of the probability weighting function. *Management Science*, 42(12):1676–1690.

Yamai, Y. and Yoshiba, T. (2005). Value-at-risk versus expected shortfall: A practical perspective. *Journal of Banking & Finance*, 29(4):997–1015.

Yang, X. (2006). Improving portfolio efficiency: A genetic algorithm approach. *Computational Economics*, 28(1):1–14.

Yiu, K.-F. C. (2004). Optimal portfolios under a value-at-risk constraint. *Journal of Economic Dynamics and Control*, 28(7):1317–1334.

Yun, Y., Gen, M., and Seo, S. (2003). Various hybrid methods based on genetic algorithm with fuzzy logic controller. *Journal of Intelligent Manufacturing*, 14(3):401–419.

Zhang, J., Chung, H. S.-H., and Lo, W.-L. (2007). Clustering-based adaptive crossover and mutation probabilities for genetic algorithms. *IEEE Transactions on Evolutionary Computation*, 11(3):326–335.