School of Electric Engineering, Computing and Mathematics

Robust and Multi-objective Portfolio Selection

Lin Jiang

This thesis is presented for the Degree of
Doctor of Philosophy
of
Curtin University

June 2020

Declaration

To the best of my knowledge and belief this thesis contains no material previously published by any other person except where due acknowledgment has been made.

This thesis contains no material which has been accepted for the award of any other degree or diploma in any university.

....

Lin Jiang

June 2020

Abstract

Portfolio selection is to allocate resources into assets. Markowitz's seminal work on mean-variance model provided the first quantitative treatment of the trade-off between investment and risk. After Markowitz's seminal work, there are tremendous amount of research on portfolio selection from both model and computational algorithms to make model portfolio theory more practical. In the Markowitz's model, risk is measured by variance which is suffered from drawbacks if distribution is not symmetry. To overcome this drawback, some more risk measurements are developed, such as semi-variance, Value-at-Risk (VaR) and Conditional-Value-at-Risk(CVaR). In the standard mean-variance model, only one-off decision is made at the beginning of the period which is maintained until the end of this period. However, investors always like to adjust their investment according to the real performance of the portfolios. Thus, multi-period portfolio selection problem has attracted extensive research interests. During portfolio selection modelling, the investment return and the variance of the portfolios are usually estimated through the historical data. However, these estimates are inexact and suffer from uncertainty. Thus, robust portfolio selection problem is considered. Although portfolio selection problem has been studied about seventy years since Markowitz's seminal work, there are still many problems unresolved due to the complex nature of the portfolio selection. In this thesis, robust and multi-objective portfolio selection problem will be further studied. New models and computational algorithms will be developed to solve the proposed models.

In Chapter 1, we will briefly introduce the portfolio selection problem and the corresponding models.

In Chapter 2, the results on existing portfolio selection are reviewed.

In Chapter 3, the distributionally robust multi-period portfolio selection problem subject to bankruptcy constraints is studied. Distributionally robustness means that the worst performance of the portfolios in terms of distribution will be optimised. For this model, we consider two cases: one is that the moment information is exactly known and the other one is the the moment information is uncertain, but within an elliptical set. For the two cases, we transform them into second-order-cone programming problems which can be easily solved by existing convex optimization toolbox. Numerical experiments are presented to illustrate our methods.

In Chapter 4, robust multi-period and multi-objective portfolio selection problem subject to no-shorting constraints and transaction costs is studied. In this model, we suppose

that the mean and variance of the investment return vector are within an elliptical data set. Then, the worst investment return and risk in the uncertainty set are optimised. For the original minimax optimisation problem, we can prove that it is equivalent to an minimax optimisation problem where the inner maximisation is one concave and one dimension. So the inner maximisation can be analytically solved. Through weighting method, we transform the original multi-objective optimisation problem into a single-objective optimisation problem which can be easily solved. Numerical experiments are presented to show the impact of the parameters' uncertainty to the performance.

In Chapter 5, we develop a nonlinear scalarisation method to solve a tri-objective portfolio selection problem. In this problem, the investment return, risk and skewness are optimised. Different from traditional linear weighting method, we develop a nonlinear scalarisation method to solve this problem. Due to the objective of skewness is non-convex, the nonlinear scalarisation method can achieve better Pareto-front solutions. Numerical experiments on the performance of nonlinear scalarisation method for different benchmarks and the tri-objective portfolio selection problem is presented.

Chapter 6 concludes the thesis and give some future directions on research of portfolio selection.

Acknowledgements

I would like to take this opportunity to express my gratitude and appreciation to those who have provided me support, advice and encouragement to complete this thesis. Without their kind assistance, it might be impossible to finish this thesis on time.

First and foremost, I would like to express my sincere gratitude to my supervisor Prof. Song Wang for the continuous support of my Ph.D study and related research, for his patience, insightful guidance and professional comments. The thesis will not be possible without his invaluable comments.

Second, I would like to specially express my gratitude to my co-supervisor A/Prof. Guanglu Zhou. His encouragement, instrument and generous sharing of his profound knowledge have made my PhD journey much more smoother and easier.

Third, I would like to specially thank Prof. Yonghong Wu for his instruments in my PhD APA scholarship application. It would be impossible without this scholarship support. In addition, I also would like to thank Dr. Qiang Long and Chuanxin Zhao, who have helped me on problem formulation and programming.

Finally, I would like to thank my family and friends for their consistent support, encouragement throughout my entire PhD study.

List of Publications

The following papers were completed during PhD candidature and have been accepted for publication:

- Chuanxin Zhao, Lin Jiang, Kok Lay Teo, A hybrid chaos firefly algorithm for three-dimensional irregular packing problem. Journal of Industrial and Management Optimization, 2020, 16(1), 409-329.
- Lin Jiang, Song Wang, Robust multi-period and multi-objective portfolio selection, Journal of Industrial and Management Optimization, DOI: 10.3934/jimo.2019130, published online.
- Qiang Long, Lin Jiang, Guoquan Li, A nonlinear scalarization method for multi-objective optimization problems, Pacific Journal of Optimization, 2020, 16(1), 39-65.

The following paper is completed during PhD candidate and currently under review:

- Lin Jiang, Song Wang, Changzhi Wu, Distributionally robust multi-period portfolio selection subject to bankruptcy constraints.

Contents

List of Tables **1**

List of Figures **2**

1 Introduction **4**

1.1 Background 4
1.2 Research Objectives 7
1.3 Thesis Organization 9

2 Literature Review on Portfolio Selection **10**

2.1 Single-Period Portfolio Selection 10
2.2 Multi-Period Portfolio Selection 13
2.3 Multi-Objective Portfolio Selection 15

3 Distributionally Robust Multi-Period Portfolio Selection Subject to Bankruptcy Constraints **18**

3.1 Introduction 18
3.2 Multi-period portfolio selection with bankruptcy constraints 20
3.3 Deterministic Tractable Reformulation and Computation 22
3.4 Numerical studies 28
3.5 Conclusion 33

4 Robust Multi-Period and Multi-Objective Portfolio Selection **34**

4.1 Introduction 34
4.2 Problem Statement 36
4.3 Problem Transformation 38
4.4 Solution strategy 40
4.4.1 Weighted-sum approach 40
4.4.2 Sub-optimization solution 41
4.5 Numerical experiments 42
4.5.1 Convex of $g(\kappa)$ 42
4.5.2 Impact of the input parameters 43
4.5.3 Numerical solutions with different parameters 45

4.5.4 Pareto-front analysis 45
4.6 Conclusion 48

5 A Nonlinear Scalarization Method for Multi-objective Optimization Problems and Applications for Portfolio Selection 50
5.1 Introduction 50
5.2 Preliminaries 51
5.3 Extended weighted sum method 53
5.4 Nonlinear scalarization method 56
5.5 Implementation 60
5.5.1 Generating λ in Algorithm EWSM 60
5.5.2 Generating θ in Algorithm NSM and SNSM 61
5.5.3 Global optimization solver 63
5.6 Mean-Variance-Skewness Portfolio Selection 63
5.7 Numerical experiments 64
5.7.1 Illustrative examples 64
5.7.2 Comparison with MOEA/D and NSGAII 66
5.7.3 Numerical test and comparisons using CEC'09 69
5.7.4 Solving the mean-variance-skewness model 73
5.8 Conclusion 73

6 Conclusions and Future Research 82
6.1 Main Contributions of this Thesis 82
6.2 Future Research Directions 83

Bibliography 84

List of Tables

4.1 Solution with five period $\varsigma = 0.001, \delta = 0.01$ 47
4.2 Solution with five period $\varsigma = 0.01, \delta = 0.01$ 47
4.3 Solution with five period $\varsigma = 0.01, \delta = 0.1$ 48

5.1 Multi-objective test problems. 64
5.2 The numerical performance evaluated by IGD. 72

List of Figures

1.1 Multi-period portfolio selection 7

3.1 The optimal $\mathbb{E}(x_t)$ with $\underline{\mathrm{x}}= 1.15 and \underline{\mathrm{x}}= 1.196$ 29
3.2 The optimal $u_2(t)$ with $\underline{\mathrm{x}}= 1.15 and \underline{\mathrm{x}}= 1.196$ 30
3.3 The optimal $u_3(t)$ with $\underline{\mathrm{x}}= 1.15 and \underline{\mathrm{x}}= 1.196$ 30
3.4 The optimal $\mathbf{u}_t$ with $\gamma_1 = 0.0001$ and $\gamma_2 = 1.2$ 31
3.5 The optimal investment return under different γ_2 32
3.6 The optimal investment return under different γ_1 32

4.1 The variation of $g(\kappa)$ in terms of κ when $\lambda = 0.1$ 43
4.2 The variation of $g(\kappa)$ in terms of κ when $\lambda = 0.5$ 44
4.3 The variation of $g(\kappa)$ in terms of κ when $\lambda = 0.9$ 44
4.4 The variation of $f_\lambda(\Delta w_t, \kappa^*)$ in terms of δ when $\lambda = 0.1$ 45
4.5 The variation of $f_\lambda(\Delta w_t, \kappa^*)$ in terms of δ when $\lambda = 0.5$ 46
4.6 The variation of $f_\lambda(\Delta w_t, \kappa^*)$ in terms of δ when $\lambda = 0.9$ 46
4.7 The variation of $f_\lambda(\Delta w_t, \kappa^*)$ in terms of ς when $\lambda = 0.5$ 47
4.8 The Pareton front with $\delta = 1$ 48
4.9 The Pareton front with $\delta = 5$ 49
4.10 The Pareton front with $\delta = 10$ 49

5.1 Geometrical meaning of linear and nonlinear scalarization methods. 56
5.2 Generating weights using systematic method. 61
5.3 Generating referential points θ for Algorithm NSM. 62
5.4 Generating referential points θ for Algorithm SNSM. 62
5.5 Objective function value set of Problems SCH, FON and KUR. 65
5.6 Solving Problem SCH using Algorithm 1. 66
5.7 Problem FON solved by NSM. 66
5.8 Problem FON solved by SNSM. 67
5.9 Problem FON solved by SNSM with Pareto sorting. 67
5.10 Problems KUR solved by NSM and SNSM. 68
5.11 Numerical performance for 2-objective problems. 75
5.12 Numerical performance for 3-objective problems. 76

5.13 Comparison respect to time consumption and number of function value evaluation. 76
5.14 Objective function value space for test problems. 77
5.15 Real Pareto frontiers for test problems. 77
5.16 Solving Problem 1 and 3 using SNSM. 78
5.17 Solving Problem 4 and 6 using SNSM. 78
5.18 Solving Problem 7 and 9 using SNSM. 78
5.19 Return-Risk Pareto-front. 79
5.20 Return-Risk Pareto front. 80
5.21 The Pareto-frontiers obtained by SNSM, MOEAD and NSGAIILS. 81

CHAPTER 1

Introduction

1.1 Background

A portfolio is a grouping of financial assets, such as stocks, bonds, commodities, currencies, asset-backed securities, real estate certificates and bank deposit. Portfolios are held by investors and/or managed by financial professionals and money mangers. Investor should construct an investment portfolio in accordance with their investment return, risk tolerance, asset diversification, etc.

An investment portfolio is to allocate the resources into different assets for the purpose of maximizing benefit while minimizing the risk or maintaining the risk to be under control. Portfolio selection is to optimally allocate investors' capital to a number of candidate securities. The process of selecting a portfolio can be divided into two stages [55]. The first stage is to estimate the future performance of a candidate portfolio through its past performances. In the second stage, the candidate portfolios will be selected based on their estimated future performances and investors' preference on return and risk.

Formally, the portfolio selection can be formulated as follows: Given a set of N assets which we may invest, we need a strategy to divide the resources among these assets, such that after a specified period of time T, the return on investment can be achieved as high as possible while minimizing the risk or maintaining the risk under a given level.

The fundamental breakthrough of solving this problem dates back to Markowitz's seminal work in 1952 [55]. In Markowitz work, this problem was formulated as a mean-variance optimization problem where the risk is measured through the variance of the candidate portfolios. In this model, an investor regards expected return as desirable and variation of return as undesirable. Let r_i be the random variables which are the future rate of return for the asset i, $i = 1, \cdots, N$, and define $z = [r_1, r_2, \cdots, r_N]^T$ which is the collection of all the random variables r_i. Denote $\mu_i = E(r_i)$, $m = [\mu_1, \mu_2, \cdots, \mu_N]^T$ and the covariance matrix $\Sigma = cov(z)$, where $E(r_i)$ means the expectation of the random variable r_i, $cov(z)$ is the covariance of the random vector z. Suppose that $[w_1, w_2, \cdots, w_N]^T$ is a set of weights which are corresponding to the investment percentages to the assets. Then, the Markowitz mean-variance model is as below:

Markowitz Mean-Variance Model:

$$\begin{aligned} \min \quad & \frac{1}{2} w^T \Sigma w \\ \text{subject to} \quad & m^T w \geq \mu_b, \\ & e^T w = 1, \ w \geq 0, \end{aligned} \tag{1.1}$$

where μ_b is the given expectation return, $e = [1, \cdots, 1]^T$.

An alternative formulation, which explicitly trades off risk and the return in the objective function through method, is as follows:

$$\begin{aligned} \max \quad & m^T w - \lambda w^T \Sigma w \\ & e^T w = 1, \ w \geq 0, \end{aligned} \tag{1.2}$$

where λ is the given weight to trade off risk and investment return.

Similarly, one can also consider to maximization return while keeping variance under a given level:

$$\begin{aligned} \max \quad & m^T w \\ \text{subject to} \quad & w^T \Sigma w \leq \sigma_p, \\ & e^T w = 1, \ w \geq 0, \end{aligned} \tag{1.3}$$

where σ_p is a given risk level.

In Markowitz mean-variance model, the market is considered without transaction cost, the short selling is not allowed, and the assets are considered to be able traded with any non-negative fractions. Variance as a risk measure has been criticized because of its symmetrical treatment of both upside and downside deviations from the mean as risk, which cannot be justified, especially, for skewed distributions. [101]. To overcome this drawback, some other risk measures are proposed.

In order to consider special aversion to returns below the mean value, downside risk measure such as the semivariance of return is introduced in the literature [76]. The semivariance is defined as the weighted sum of square deviations below this mean value [76]. Mathematically, Mean-Semivariance model for the portfolio selection can be defined as below:

Mean-Semivariance Model:

$$\begin{aligned} \min_{w} \quad & \int_{\mathbb{R}^N} \min(r^T w - \mu_b, 0)^2 dP \\ \text{subject to} \quad & m^T w \geq \mu_b, \end{aligned}$$

$$e^T w = 1,\ w \geq 0, \tag{1.4}$$

where P is the joint distribution of r.

In addition to semi-variance, there are some other downside risk measures, such as Value-at-Risk (VaR) and Conditional-Value-at-Risk (CVaR), which are widely used in the literature to describe the risk. VaR measures the maximum likely loss of a portfolio from market risk with a give confidence level $(1-\alpha)$. For example, if VaR is valued as 10,000 with 95% confidence level, it means that there is only a 5% chance that the loss will be greater than 10,000. The higher the confidence level, the less the chances the loss is out of the value. [54]. Mathematically, VaR at a given confidence level $1-\alpha$ is the maximum expected loss that the portfolio cannot exceed with probability α,

$$\mathrm{VaR}_\alpha(w) = \min\left\{\zeta \in \mathbb{R} : \Psi(w,\zeta) \geq \alpha\right\}, \tag{1.5}$$

where $\Psi(w,\zeta)$ is the probability of the loss not exceeding a threshold ζ. Suppose that the probability function is P, then

$$\Psi(w,\zeta) = \int_{-r^T w \leq \zeta} dP. \tag{1.6}$$

Then, the mean-VaR model can be stated as below:

Mean-VaR Model:

$$\begin{aligned} \min_w \quad & \mathrm{VaR}_\alpha(w) \\ \text{subject to} \quad & m^T w \geq \mu_b, \\ & e^T w = 1,\ w \geq 0, \end{aligned} \tag{1.7}$$

As a measure of risk, VaR has its limitations, such as lacking subadditivity, convexity, and not coherent [88]. An alternative risk measurement, CVaR, is coherent with attractive properties including convexity. Thus, the model based on CVaR is easier than VaR to compute from the mathematical perspective. Mathematically, CVaR is defined as the conditional expectation of the portfolio loss exceeding or equal to VaR [68]:

$$\mathrm{CVaR}_\alpha(w) = \frac{1}{1-\alpha}\int_{-r^T w \geq \mathrm{VaR}_\alpha(x)} -r^T w dP \tag{1.8}$$

Mean-CVaR Model:

$$\begin{aligned} \min_w \quad & \mathrm{CVaR}_\alpha(w) \\ \text{subject to} \quad & m^T w \geq \mu_b, \end{aligned}$$

$$e^T w = 1, \ w \geq 0, \tag{1.9}$$

In portfolio selection problems, there are at least two-objectives: return and risk. In some applications, there have even more objectives, such as the diversification of the investment and the liquidity of the portfolios. In the above mentioned models, either only one objective is minimized and the others are put into the constraints or the objectives are weighted together as only one objective. For the first case, how to determine an exact threshold for an investor is rather challenging. For the second case, how to determine this weight is difficult as different investors have different preferences to the objectives. Thus, to present all potential solutions to the investors is much important and study the multi-objective portfolio selection are paramount to applications. In this thesis, we will develop a meta-heuristic based method for the portfolio selection.

In the above model, the portfolio selection problem has been formulated as a standard optimization problem where the risk is minimized while maintaining the expected investment return to be above a desired level. Furthermore, only a static portfolio selection is considered. In practice, investors always prefer to invest long-term assets for obtaining investment return. In this scenario, the investors are required to adjust the assets held according to the assets financial performance from time to time as shown in Figure 1.1 In Figure 1.1, $x_{i+1} = x_i + \Delta x_i$, $i = 0, 1, \cdots, T-1$. Generally, a multi-period portfolio

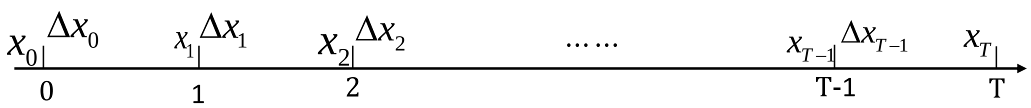


Figure 1.1: Multi-period portfolio selection

selection problem is heavily depending on the given dynamics. For different dynamics, the solution methods are different. In this thesis, we will study multi-period portfolio selection with uncertainty.

1.2 Research Objectives

The aims of the thesis are to study portfolio selection under different scenarios. In particularly, we will study the following portfolio selection problems:

(i) Multi-period portfolio selection with moment uncertainty and subject to bankruptcy;

(ii) A nonlinear scalarisation method for multi-objective optimisation and applications in portfolio selection;

(iii) Multi-period and multi-objective portfolio selection with bound uncertainty for the return expectation and covariance matrix.

Distributionally Robust Multi-Period Portfolio Selection Subject to Bankruptcy Constraints:

In portfolio selection, we need to know the future performance of the candidate portfolios which is estimated from the historical data. Based on the historical data, we can estimate the mean and covariance of the excess return of the candidate portfolios. In some of the existing works, the excess return is assumed to follow a normal distribution. Clearly, this assumption is too strong as there are too many factors might affect the excess return of the candidate portfolios in practice. To address this shortcoming, we introduce the distributionally robust optimization to study the distribution of the excess return is unknown in advance. In addition, the estimates of the mean and covariance of the excess return of the candidate portfolios are not exact. In this thesis, we study portfolio selection with inexact estimates of the mean and covariance, but within a bound set. For the two cases, we will derive tractable algorithms to solve them.

Robust Multi-Period and Multi-Objective Portfolio Selection:

In the multi-period mean-variance model, the investment return and the risk are usually weighted together to be a single objective to optimize. However, different investors have different preferences on return and risk. Furthermore, there is lacking a unified way to determine the weight. Under this circumstance, to present all the potential solutions to the investors is important so that they can choose the one that is best suitable for them. As mentioned in the above, to consider the uncertainty of the return is important since the exact information is always not available. In the above model, the mean and variance are considered as uncertain and varied within a bound set. In this model, instead of formulating such a portfolio selection problem as distributionally robust optimization problem, we formulate it as a deterministic optimization problem where the excess return is considered to be bounded within a given set.

A nonlinear scalarisation method for multi-objective optimisation problems and applications to portfolio selection:

Portfolio selection problem is a multi-objective optimisation problem in nature. If skewness is optimised, the corresponding optimisation problem is nonconvex. The existing linear weighting method might not provide good approximations of the solutions in Pareto-front. In Chapter 5, we will develop nonlinear scalarisation method in stead of traditional linear weighting methods to transform a multi-objective optimisation problem into a single-objective optimisation problem. We test numerical performance of the proposed nonlinear scalarisation method through a wide of benchmarks. Then, this proposed method is introduced to solve a tri-objective portfolio selection problem.

1.3 Thesis Organization

This thesis is organised as follows:

- *Chapter 2*: This chapter presents a survey of existing results on portfolio selection.
- *Chapter 3*: This chapter aims to address the distribution and moment uncertainty in the multi-period portfolio selection problem. We firstly study multi-period portfolio selection problem with the given mean and covariance of the excess return but unknown distribution. Under this assumption, the original portfolio selection problem is formulated as a distributionally robust optimization problem. This problem is further transformed into an equivalent deterministic optimization problem which can be solved easily by the existing optimization software. Then, we study the case that both the estimates of the mean and the covariance of the excess return are also uncertain but within a bounded set. For such a problem, we investigate its theoretical characteristics and prove that it can also be transformed into a tractable convex optimization problem.
- *Chapter 4*: This chapter studies the multi-period multi-objective portfolio selection with the investment return uncertainty. We firstly show how to derive a bound set to bound the investment uncertainty. Then, the problem is formulated as min-max multi-objective optimisation problem. For the inner maximization problem, we prove that it can be transformed into a maximization problem with only one variable. We further prove that for a given weight, the inner maximization problem is concave and thus the optimal solution is either achieved at the boundary point or in the equilibrium point within the interval if it has. Thus, the inner maximization problem is easily solved. Since there are only two objectives, the Pareto-front can be plotted against the weight.
- *Chapter 5*: This chapter develops a nonlinear scalarisation method to solve multi-objective optimisation problems. Different from traditional methods to transform a multi-objective optimisation problem into a single-objective optimisation problem through linear scarlarisation, this nonlinear scalarization method is to transform a multi-objective optimisation problem into a single-objective optimisation problem through nonlinear scalarisation. We will investigate the theoretical characteristics and numerical performance of the proposed method.
- *Chapter 6*: A brief summary of the thesis contents and its contributions are given in the final chapter. Recommendation for future works is given as well.

CHAPTER 2

Literature Review on Portfolio Selection

In this chapter, we will review the existing results in portfolio selection.

2.1 Single-Period Portfolio Selection

For an investor, the challenging problem is how to allocate their current wealth over a number of available portfolios, such as stocks, bonds and derivatives, to maximize the return while minimizing the risk. Such a problem is referred to as portfolio selection [104]. In dealing with this fundamental issue, Markowitz in his seminal work [55] proposed a mean-variance model, where the risk is measured by the variance. A practical advantage of the Markowitz model is that this problem has been formulated as a convex quadratic program, which can be solved efficiently. Due to this fundamental contribution, Harry Markowitz received the 1990 Nobel Prize in Economics. As the Swedish Academy of Sciences put it "his primary contribution consisted of developing a rigorously formulated, operational theory for portfolio selection under uncertainty" [76].

After Markowitz's seminal work [55], there are tremendous amount of research on portfolio selection from both model and computational algorithms to make model portfolio theory more practical. In the Markowitz's Mean-Variance model, the risk is measured by the variance. Under this scenario, the variability of the variance is minimized and thus, the variability of the actual return over the average return is minimized. If the return follows a normal distribution, the mean-variance model will produce an efficient strategy since the symmetry of the distribution. However, in practice, the normal distribution of the investment return is highly unlikely. To overcome this drawback, the semi-variance (or downside) risk is introduced [67] for portfolio selection. Theoretically, semi-variance model should produce a better solution since an investor only worries about underperformance, not about overperformance. However, due to the endogenous of the semicovariance matrix, the corresponding optimization problems are intractable [67] which results in the popularity of the mean-variance model. In [23], a heuristic method is introduced to solve the mean-semivariance model since its intractability. In [67], it has shown that although minimizing the semivariance is more in line with the true preferences of a rational investor, but minimizing the variance usually achieves a lower downside deviation and a higher Sortino ratio because it can be estimated more accurately.

In addition to semivariance, VaR is one of the most popular risk measures [54]. For a given confidence level and a particular time horizon, a portfolio's VaR is the maximum loss one expects to suffer at that confidence level by holding that portfolio over that time horizon [2]. Different from Mean-Variance model, the optimization of Mean-VaR model is NP-hard and thus, a global optimal solution is hard to be obtained. In [2], VaR is introduced for portfolio selection and its economic implications is studied. Through comparing with Mean-Variance model, it shows that the higher variance portfolio has less VaR. It reveals that for certain risk-averse agents, the portfolios with larger standard deviations might be selected if Mean-VaR model is adopted instead of Mean-Variance model [2]. Since Mean-VaR model is NP hard, a meta-heuristic based approach is developed in [54] to solve the model since it has multiple local extrema and discontinuities when the real-world constraints are incorporated.

An alternative risk measure to VaR is CVaR, which is also known as mean excess loss, mean shortfall or tail VaR [68]. CVaR is defined as the weighted average of VaR and losses strictly exceeding VaR for general distribution [68]. It has been shown that CVaR is a coherent risk measure [3] with many attractive properties including convexity. Thus, CVaR is a convex risk measure which is much easier to be solved than VaR. In [68], Mean-CVaR portfolio optimization problem is transformed into a linear programming (LP) problem based on the CVaR measure with discrete sample approximation. In this approximation, the dimension of LP is $n + 1 + S$, where n is the dimension of the decision variables and S is the number of the samples. To achieve good approximations, the samples should be large enough which means S is a very large number. Thus, the resulted LP problem has very high dimension and expensive to be solved even it is linear. Some approaches are proposed to address this shortcoming. In [5], a discrete gradient method is proposed to solve nonsmooth portfolio optimization problem with CVaR. Through numerical comparisons with LP-based approach, it has shown that non-smooth based optimization methods can even achieve better performance if the scenarios are large. In [81], a smoothing technique is developed to solve the scenario-based CVaR. Comparing with the nonsmooth approach, the smoothing method can use ordinary derivative conveniently and can retain nice convergence properties and the convexity for the scenario-based CVaR problem [81]. The numerical performances also show its superior to LP-based approach.

To make the portfolio selection model more practical, the transaction cost should be considered. There are different techniques to handle the transaction cost [8]. A recent technique to model the transaction cost is to include cardinality constraint in the model which means to select portfolios as sparse as possible. Since the cardinality constraint is included 0-norm in the model, the corresponding optimization problem becomes non-convex and even discontinuous. How to solve such an optimization problem with large candidate portfolios is still challenging. Currently, the available methods for the cardinality constrained portfolio optimization can be classified as two categories: exact-based method and heuristic-based method. In [25], an exact-based approach is developed to solve such a problem through exploring the special structures and geometric properties behind. By introducing the Lagrangian relaxation to the primal problem and modifying the primal objective function to some separable functions, a sharp lower bounding scheme is developed [25] which is further integrated into a branch-and-bound algorithm to solve the original problem. In [79], a completely positive programming reformulation of the cardinality constrained portfolio optimization problem is proposed as the lower bound. Through integrating a heuristic method which is used to obtain a good feasible solution, the original cardinality constrained portfolio optimization problem is solved through a Branch-and-Bound method. In addition to exact approaches, there are extensive meta-heuristic based methods to solve such an problem. In [38], an artificial bee colony algorithm with a feasibility enforcement procedure along with an infeasibility toleration procedure is proposed to solve cardinality constrained portfolio optimization problem. Infeasibility toleration procedure is introduced to allow the solution to violate bounding constraints temporarily while the repair mechanism ensured the number of assets to be held in the portfolio to stick with the desired number. In [70], a genetic-based approach is proposed to solve cardinality constrained portfolio optimization with transaction costs. The problem is first formulated as a mixed-integer quadratic problem. To conduct genetic search, the candidate portfolios are encoded using a set representation to handle the combinatorial aspect of the optimization problem. Besides specifying which assets are included in the portfolio, this representation includes attributes that encode the trading operation (sell/hold/buy) performed when the portfolio is rebalanced [70]. In [39], a hybrid approach through integrating ant colony optimization, genetic algorithm and artificial bee colony optimization is proposed to solve cardinality constrained portfolio optimization.

Another important research topic on portfolio selection is on the parameter uncertainty in the optimization model. In fact, the parameters used in the model, such as the return of the investment and the covariance matrix which are estimated through historical data, are suffered from uncertainty. If the decision made under this uncertain environment is implemented in practice, it might result in error-maximized and investment-irrelevant portfolios [60]. Many methods are developed to address this problem. In [60], it suggests resampling the mean returns and the covariance matrix within a confidence region around a nominal set of parameters is proposed, and then aggregating the portfolios obtained by solving a Mean-Variance problem for each sample. In Part V of [102], scenario-based stochastic programming models is proposed to handle the uncertainty in parameters. However, both of the two kind of methods do not provide any guarantees on the portfolio performance and become inefficient if the number of assets are large [31]. To ensure the portfolio performance, robust models based on Mean-Variance model are proposed in [31] to address parameter uncertainty and estimation errors. In the model, the uncertainty set is modeled within either a boxed set or an elliptical set. Then, the corresponding optimization problem is transformed into a second-order-cone programming (SOCP) which can be solved by the existing software efficiently. In [30], a VaR-based robust model for portfolio selection is proposed. If the moments are exactly known, the problem is equivalent to a SOP. If the moments are suffered from uncertainty which lie in a bounded set, the problem is transformed into a Semi-Definite Programming (SDP). In [100], the worst CVaR CVaR) model is proposed to against the uncertainty in portfolio selection. Mixture distribution uncertainty, box uncertain and ellipsoidal uncertainty are studied. For all the three scenarios, the corresponding CVaR can be transformed into a convex programming which can be easily to be solved.

2.2 Multi-Period Portfolio Selection

In single-period portfolio selection problem, an one-off decision at the beginning of the period is made and maintained until the end of this period. However, in the real world, investors often adjust their wealth from time to time by taking into consideration the volatile market conditions. Thus, multi-period portfolio selection has attracted extensive research interests. Seminal work on multi-period portfolio selection can be dated back to [59]. In this work, Ito's lemma and stochastic dynamic programming are introduced to analyze optimal continuous-time dynamic portfolio selection which is still being widely used. In [17], it shows that a multi-period asset allocation problem with the final value of the portfolios to be maximized can be solved as multi-stage stochastic linear programs through combination of Benders decomposition and importance sampling. However, only the return of the investment is considered without variance being considered. In [46], a dynamical mean-variance portfolio selection problem is studied. Due to the nonseparability in the sense of dynamical programming, solving the original problem directly is difficult. In [46], through embedding the problem into a tractable auxiliary problem that is separable, an analytical solution is derived for the original problem. The problem considered in [46] is without constraints. In [14], mean-field formulation is introduced to solve multi-period mean-variance portfolio selection problem. Under this framework, some simple constraints can be included to derive analytical solution.

In the above literature, the methods are focusing on non-constrained portfolio selection problems. In applications, adjustment of portfolio in the periods is always suffered from constraints, such as no-shorting, risk allowance and bankruptcy constraint. In [11], no-shorting constraint is considered for a multi-period mean-variance portfolio selection model. It is shown that the optimal portfolio policy is piecewise linear with respect to the current wealth level. Then, semi-analytical expression of the piecewise quadratic value function is derived. In [26], a dynamic mean-variance portfolio selection problem with the time cardinality constraint and correlated returns is studied. Its analytical expression of the efficient mean-variance frontier is derived through embedding scheme. In [12], a multi-period mean-variance portfolio selection with management fees is studied. This problem is reformulated as a multi-period mean-variance portfolio selection problem with no-shorting constraint. Based on this, its semi-analytical optimal policy which is a linear threshold-type policy is derived based on the methods in [11]. In [89], stochastic interest rate is considered in a multi-period mean-variance portfolio selection problem. Through invoking dynamic programming and the Lagrangian duality, analytical solution of the original problem is also established.

The above literature is mainly focusing on the multi-period mean-variance model. In addition to the variance, CVaR and VaR are two widely used risk measures as mentioned in the above section. CVaR as a risk measure in the multi-period case is time-inconsistency [4] which prevents direction application of conventional stochastic control to multi-period portfolio selection. In [15], multi-period mean-CVaR portfolio selection problem is discussed. Due to the inconsistency of the model, the truncated global optimal policy is not optimal for the remaining short term problem. The pre-committed policy is obtained through linear programming and time consistency policy is derived through by solving a series of integer programming problems in [15]. In [77], this problem is further discussed. Through embedding the original, time inconsistent problem into a family of time-consistent expected utility maximization problems with a piecewise linear utility function, the multi-period mean-CVaR portfolio selection problem can be analytically solved. It has shown that the optimal investment strategy is a fully adaptive feedback policy and a cumulated amount invested in the risky assets is of a characteristic V-shaped pattern as a function of the current wealth [77]. The above derived optimal mean-CVaR investment strategies is based on pre-committed strategies. However, as investors would like to choose the local optimal strategy intermediate time instants, it is hard to maintain that strategy although it has better global investment performance in terms of the initial time in practice [13]. How to balance the conflicts between an investor's global and local investment interests is important. In [13], tractable computational methods are proposed to coordinate this conflicted interests. In addition to CVaR as a risk measure, VaR is considered as a risk measure in multi-period portfolio selection. In [61], a multi-period mean-VaR model for portfolio selection is studied. Genetic-based algorithm is proposed to solve such a problem.

In addition to variance, CVaR and VaR as risk measures, the lower partial moment is also widely used in multi-period portfolio selection. In [49], a robust multi-period portfolio selection model based on downside risk with asymmetrically distributed uncertainty set is studied. Through introducing distributionally robust, the original model can be solved through solving a second-order cone optimization problem. In [87], multi-period semi-variance portfolio selection problem is studied. The downside risk based on the semi-variance is introduced in the model. Due to the non-smoothness of the objective function caused by semi-variance, a hybrid genetic algorithm is proposed to solve the problem. In [62], multi-period portfolio selection is formulated as minimizing one-sided deviation from a target wealth level and maximizing the expected end-of-horizon wealth problem. The trade-off between two objectives is controlled by a given weight. Through introducing a piece-wise linear penalty function, the original problem can be solved through solving a series of linear programming problems. In [45], different most of the methods where the return rates of the candidate portfolios are considered as determined, a multi-period portfolio selection problem under uncertain investment returns with bankruptcy constraint is considered. The proposed uncertain optimization problem is transformed into the crisp optimization models and then a genetic algorithm is proposed to solve the problem. In [40], instead of using the average standard deviation of portfolio for all portfolios, the estimation of standard deviation of itself is used to calibrate the risk. The kernel distribution function is introduced to calibrate the distribution of the random variables.

2.3 Multi-Objective Portfolio Selection

Portfolio selection is to allocate resources to a number of portfolios. Based on their preference, investors need to trade-off investment return and risk during the resource allocation process. Due to the intrinsic multi-objective nature of the problem, evolutionary-based multi-objective algorithms have been widely developed to solve this class of optimization problems. In [47], a cardinality constrained multi-objective portfolio selection problem is studied where return is maximized and the risk is minimized. The risk is measured by variance. In addition to the cardinality constraint, floor and ceiling constraints are also imposed. The floor constraint is to refrain very small weighting of any candidate portfolios to be selected. A multi-objective evolutionary algorithm through customized mutation and recombination is designed to solve the problem. The performance of the algorithm has been tested through a wide range of data sets which size is varying from 31 to 1317. In [73], the mean-variance multi-objective portfolio selection problem is further discussed. Additional constraints, such as round-lot constraints and pre-assignment constraints, are taken into account. The round-lot constraints is to restrict some candidate portfolios can only be invested in multiples of a certain amount, while the pre-assignment constraints is to include some candidate portfolios which must be invested. A particle swarm algorithm with an adaptive ranking is introduced to solve the problem. In [58], the transaction cost is included in multi-objective portfolio selection. Three models are considered: 1) mean, variance and transaction costs as objectives; 2) mean, VaR and transactions as objectives; 3) mean, CVaR and transactions as objectives. In all the three models, cardinality constraint, quantity, pre-assignment, self-financing constraint and the equality constraints arising due to the consideration of transaction cost are incorporated in the model. An evolutionary-based algorithm is proposed where the equality constraint is handled through a repair algorithm. In [64], mean-variance portfolio selection problem is considered as a bi-objective optimization problem. This bi-objective optimization problem has been transformed into a single objective optimisation problem through fuzzy normalization. Then, invasive weed optimisation algorithm is introduced to solve this problem. In [57], a multi-objective portfolio selection problem with cardinality, pre-assignment, budget, quantity and round-lot constraints. A evolutionary-based multi-objective optimisation algorithm is proposed to solve the model. In [65], portfolio selection is formulated as a bi-objective optimisation problem. A hybrid bi-objective algorithm combining with the respective advantages of local search algorithm, evolutionary algorithm and QP with a pre-selection strategy is developed to solve the problem. In [63], mean-variance-skewness model for multi-objective portfolio selection is studied and a particle swarm optimisation algorithm is proposed to solve the problem.

There are also many results available to address the uncertain nature in portfolio selection. In [44], a joint probability constraint with random right-hand side vector is included in multi-objective portfolio selection problems. Mixed-integer linear programs are used to reformulate and approximate the original multi-objective probabilistically constrained programs. In [32], to improve the reliability of the investment return and risk measurement, a robust optimisation for a multi-objective product portfolio problem is studied. The future demand of each product and the risk index of each product are calculated through an neural network based algorithm. In [36], possibilistic mean value and variance of continuous distribution are introduced for multi-objective portfolio selection rather than probability distributions. A trapezoidal possibility distribution as the possibility distribution of the rates of returns on the securities is used to transform the original fuzzy optimisation problem into a deterministic one. In [72], a multi-objective robust possibilistic model for technology portfolio optimisation problem is studied. The risk and the created jobs as two objectives to be optimized and the multi-objective robust possibilistic programming approach is introduced to handle the parameter uncertainty. In [1], a multi-objective portfolio selection problem is modeled as a stochastic programming where the uncertain parameters are supposed to follow normal distribution. Then, goal programming and compromise programming are introduced to solve this multi-objective optimisation problem. In [29], multi-objective portfolio selection problem is formulated as a robust goal programming. The parameter uncertainty is handled through robust optimisation which allows the uncertain parameter to take values according to a symmetric distribution with a mean equal to the nominal values. In [37], multi-period and tri-objective uncertain portfolio selection problem is studied where the asset returns are considered as uncertain variables. The augmented weighted Tchebycheff program is introduced to transform the original tri-objective optimisation problem into a single objective optimisation problem. A particle swarm optimisation algorithm is designed to solve the problem.

CHAPTER 3

Distributionally Robust Multi-Period Portfolio Selection Subject to Bankruptcy Constraints

3.1 Introduction

Mean-variance portfolio selection was proposed by Markowitz in his seminal work [55] which was for single-period investment model. In this model, the return of investment measured by the mean of wealth is maximized while the risk measured by variance of portfolios to be selected is minimized. This model is then extended to the multi-period case [46]. In a multi-period mean-variance model, if the constraints are simple, this problem can be solved analytically. For example, analytical optimal portfolio policy and analytical expression of the mean-variance efficient frontier were derived in [46] through introducing an auxiliary parametric formulations which was to overcome nonseparable of the original problem in the sense of dynamic programming. Based on this technique, several multi-period portfolio selection problems are discussed, including the case with no shorting constraints [11], stochastic interest rate [89]. Instead of using auxiliary parametric formulations to tackle the issue of non-separability, a mean-field framework is introduced to tackle directly the issue of non-separability and derive the optimal policies analytically in [14]. Unfortunately, the analytical solution for a multi-period mean-variance model can be obtained only for those structurally simple cases.

Variance as a risk measure has been widely criticized by practitioners as it equally weights desirable positive returns and undesirable negative ones [54]. To circumvent this drawback, the semi-variance risk measure which only measures the variability of returns below the mean is introduced to replace variance [49]. Another typical kind of risk measures are Value-at-Risk (VaR) and Conditional-Value-at-Risk (CVaR). VaR is the quantile of the loss at a specified confidence level while CVaR is the conditional expected value of loss exceeding the VaR [13]. However, as a measure of risk, VaR lacks sub-additivity and convexity which leads the corresponding portfolio selection problem to being non-convexity [7]. To overcome the shortcomings of VaR, CVaR is proposed by Rockfellar and Uryasev in [68] which is proved to be coherent and convex. VaR and CVaR have been widely used in portfolio selection for both single-period and multi-period cases [13, 54]. In [54], single-period portfolio selection with mean-VaR model is studied which is a non-convex NP-hard optimization problem and an evolutionary based algorithm is proposed to solve the problem. In [88], a non-parametric approach is introduced to estimate the density of the loss function which leads to a convex formulation of the original portfolio selection problem. In [13], time inconsistency in multi-period mean-CVaR model is studied and time-consistent and self-coordination strategies are proposed. The proposed time-consistent strategy is a piecewise linear function of the wealth level with parameters which can be obtained through solving a series of mixed-integer programming problems off line. The self-coordination strategy is formulated as a convex program with a quadratic constraint.

In practical investment, the investment return is highly relying on the market. The underlying return distribution parameters, such as its expectation and covariance, cannot be obtained exactly in advance. In the recent years, there is a large amount of work to address this issue which is called as distributionally robust optimization [20]. Distributionally robust optimization is to handle the distributional uncertainty in stochastic optimization problems where the worst case of objective function and/or constraints are optimized under the given moment information. This model has been widely applied in portfolio selection to estimate the worst case of the investment risk. For example, in [30], for the given bounds on the mean and covariance matrix of returns, the worst-case VaR model of portfolio selection problem is studied. Through duality analysis, the proposed model can be cast as a semi-definite programming (SDP) which can be easily solved by existing convex optimization tools. In [100], the worst-case CVaR in robust portfolio selection is studied. For some special cases, it has been proved that the original robust optimization problem can be transformed into an equivalent SOCP. In [20], a distributionally robust optimization problem with uncertain moment information is studied. Through a series of duality analysis, the original problem can still be cast as a SDP although the mean and covariance are both suffered from uncertainty. Numerical experiments show that the daily return obtained under this distributionally robust framework is more reliable while not sacrificing daily utility. In [10], tight bounds on the expected values of several risk measures are studied. It has been shown that a single-period portfolio selection problem without additional constraints can always be solved analytically if the disutility function is in the form of lower partial moments (LPM), VaR or CVaR. For the multi-period portfolio selection problem. For the multi-period optimization problem, it has been shown that the problem can still be solved analytically in [51]. In [49], a robust multi-period portfolio selection under downside risk LPM with asymmetrically distributed uncertainty set is studied. A computationally tractable approximation approach is proposed to solve the original problem.

For existing multi-period portfolio selection problems, most of them have ignored constraints and thus, the original problem can be solved analytically through applying dynamical programming. However, investment portfolios are always required to satisfy various constraints which are determined by the investment strategy. For example, chance constraints are considered in [78] where the chance constraints are handled by the one-sided Chebyshevs inequality. Clearly, this approximation is not tight and thus the optimality of the obtained solution cannot be guaranteed. To overcome this problem, we will study multi-period portfolio selection under distributionally robust framework. In our discussions, we will study both the cases where the mean and covariance of returns are either known exactly or suffered from uncertainty. We will show that this problem can also be transformed into an equivalent SOCP which can be easily solved. Then, some numerical experiments are presented to illustrate and compare our proposed methods with those existing.

3.2 Multi-period portfolio selection with bankruptcy constraints

We consider a financial market with $n+1$ assets available to be invested which consist of one cash riskless asset and n risky assets in a time horizon with T time periods. Let the cash riskless asset be labeled as 0 and n risky assets be labeled as $1, \cdots, n$. Here the time period can be any time unit in accordance with real applications. Let s_t be the deterministic return of the riskless asset at period t and e_t^i be the random return of the risky asset i, $i = 1, \cdots, n$. Denote the vector $\mathbf{e}_t = [e_t^1, \cdots, e_t^n]^T$ to be the collection of all risky returns and the excess return vector of risky assets $\mathbf{p}_t = [p_t^1, \cdots, p_t^n] = [e_t^1 - s_t, \cdots, e_t^n - s_t]^T$. In the following discussions, we assume that the vector $\mathbf{e}_t, t = 0, 1, \cdots, T$, are statistically independent and the only information known is its first two unconditional moments, i.e., its mean $\mathbb{E}(\mathbf{e}_t) = [\mathbb{E}(e_t^1), \cdots, \mathbb{E}(e_t^n)]^T$ and its $n \times n$ positive definite covariance $\mathrm{Cov}(\mathbf{e}_t) = \mathbb{E}(\mathbf{e}_t \mathbf{e}_t^T) - \mathbb{E}(\mathbf{e}_t)\mathbb{E}(\mathbf{e}_t^T)$.

We suppose that an investor enters the market at the initial time period $t = 0$ with the wealth x_0. The investor allocates x_0 among the riskless asset and the n risky assets at the beginning of period 0 and reallocates the wealth at the beginning of each of the following period. Let x_t be the wealth of the investor at the beginning of period t and u_t^i be the amount allocated to the $i-th$ risky aseet, $i = 1, 2, \cdots, n$, at the period t, $t = 1, 2, \cdots, T-1$. We suppose that there is no transaction cost or tax to be charged during wealth reallocations. Then, the dynamics of the wealth follows the following stochastic process:

$$\begin{aligned} x_{t+1} &= \sum_{i=1}^{n} e_t^i u_t^i + \left(x_t - \sum_{i=1}^{n} u_t^i \right) s_t \\ &= s_t x_t + \mathbf{p}_t^T \mathbf{u}_t, \ \ t = 0, 1, \cdots, T-1. \end{aligned} \tag{3.1}$$

If we use probability to measure the risk and seek to maximize the terminal wealth, a multi-period portfolio selection problem with bankruptcy constraints can be formulated as the following:

$$\begin{aligned} \text{MPS:} \quad & \max \mathbb{E}(x_T) \\ \text{s.t.} \quad & x_{t+1} = s_t x_t + \mathbf{p}_t^T \mathbf{u}_t \\ & \mathbb{P}_t(x_t \geq \underline{x}) \geq 1 - \varepsilon, \ \ t = 1, \cdots, T, \end{aligned} \tag{3.2}$$

where $\underline{x}$ is the disaster level, ε is a constant to show the acceptable maximum probability of bankruptcy set by the investor and $\mathbb{P}_t$ means the probability under the distribution $\mathbb{P}_t$, $t = 1, \cdots, T$. To avoid shorting selling and maintain self-finance, the following constraints are appended:

$$\mathbb{P}_t \left(x_t - \sum_{i=1}^{n} u_t^i \geq 0 \right) \geq 1 - \varepsilon, \ \ t = 1, \cdots, T-1. \tag{3.3}$$

In practice, to get the exact distribution of $\mathbb{P}_t$ is impossible. To overcome this difficulty, most of existing results are replacing the probability constraints (3.2) by standard constraints through using Tchebycheff inequality [14]. Although this approximation leads to a easily solved problem, the solution obtained is usually too conservative. Different from current methods, we will formulate this problem as a robust optimization problem. Let $\bar{\boldsymbol{\mu}}_t$ and $\bar{\boldsymbol{\Sigma}}_t$ be the estimates of the mean and covariance of the random vector $\mathbf{P}_t$ based on the historical data. If these estimates are accurate, we define

$$\mathcal{P}_t^1 = \left\{ \mathbb{P}_t \in \mathcal{M} : \mathbb{E}_{\mathbb{P}_t}(\mathbf{p}_t) = \bar{\boldsymbol{\mu}}_t, \ \mathbb{E}_{\mathbb{P}_t}\left[(\mathbf{p}_t - \boldsymbol{\mu}_t)(\mathbf{p}_t - \boldsymbol{\mu}_t)^T \right] = \bar{\boldsymbol{\Sigma}}_t, t = 1, \cdots, T \right\}. \tag{3.4}$$

where $\mathcal{M}$ is the set of all probability distributions, $\mathbb{E}_{\mathbb{P}_t}(\cdot)$ means the expectation under the distribution $\mathbb{P}_t$. However, in practice, these estimates are usually inadequate. To incorporate the uncertainty of estimates, we consider the following uncertainty set

$$\mathcal{P}_t^2 = \left\{ \mathbb{P}_t \in \mathcal{M} : \begin{array}{l} (\mathbb{E}_{\mathbb{P}_t}(\mathbf{p}_t) - \bar{\boldsymbol{\mu}}_t)^T \bar{\boldsymbol{\Sigma}}_t^{-1} (\mathbb{E}_{\mathbb{P}_t}(\mathbf{p}_t) - \bar{\boldsymbol{\mu}}_t) \leq \gamma_1, \\ \mathbb{E}_{\mathbb{P}_t}\left[(\mathbf{p}_t - \bar{\boldsymbol{\mu}}_t)(\mathbf{p}_t - \bar{\boldsymbol{\mu}}_t)^T \right] \leq \gamma_2 \bar{\boldsymbol{\Sigma}}_t, \end{array} t = 1, \cdots, T. \right\} \tag{3.5}$$

The first constraint in (3.5) describes how the estimate $\bar{\boldsymbol{\mu}}_t$ is close to $\mathbb{E}_{\mathbb{P}_t}(\mathbf{p}_t)$ while the second constraint in (3.5) enforces the covariance estimate to be bound in a semidefinite cone defined by a matrix inequality. Instead of replacing the probability constraints by the inequalities obtained through using the Tychebycheff inequality, we consider the following distributionally robust portfolios selection model:

$$\begin{aligned}
\text{DRMPS:} \quad & \max_{\mathbf{u}} \min_{\mathbb{P}_t \in \mathcal{P}_t} \; \mathbb{E}(x_T) \\
\text{s.t.} \quad & x_{t+1} = s_t x_t + \mathbf{p}_t^T \mathbf{u}_t \\
& \min_{\mathbb{P}_t \in \mathcal{P}_t} \mathbb{P}_t(x_t \geq \underline{x}) \geq 1 - \varepsilon, \;\; t = 1, \cdots, T \qquad (3.6) \\
& \min_{\mathbb{P}_t \in \mathcal{P}_t} \mathbb{P}_t \left(x_t - \sum_{i=1}^{n} u_t^i \geq 0 \right) \geq 1 - \varepsilon, \;\; t = 1, \cdots, T-1, \qquad (3.7)
\end{aligned}$$

where $\mathcal{P}_t$ can be either $\mathcal{P}_t^1$ or $\mathcal{P}_t^2$.

3.3 Deterministic Tractable Reformulation and Computation

In this part, we will reformulate Problem DRMPS as an equivalent deterministic problem without chance constraints, which is computationally tractable. In light of (3.1) and the independence of $\mathbf{p}_t$, $t = 1, \cdots, T$, we have

$$\begin{aligned}
\mathbb{E}(x_t) &= T_0^t x_0 + \sum_{i=1}^{t-1} T_i^t \boldsymbol{\mu}_{i-1}^T \mathbf{u}_{i-1} + \boldsymbol{\mu}_{t-1}^T \mathbf{u}_{t-1}, \qquad (3.8) \\
\mathrm{Var}(x_t) &= \sum_{i=1}^{t-1} (T_i^t)^2 \mathbf{u}_{i-1}^T \bar{\boldsymbol{\Sigma}}_{i-1} \mathbf{u}_{i-1} + \mathbf{u}_{t-1}^T \bar{\boldsymbol{\Sigma}}_{t-1} \mathbf{u}_{t-1}, \qquad (3.9)
\end{aligned}$$

where $T_0^t = \prod_{j=0}^{t-1} s_j$, $T_i^t = \prod_{j=i}^{t-1} s_j$, $\mathbb{E}_{\mathbb{P}_t}(\mathbf{p}_t) = \boldsymbol{\mu}_t$ and $\mathbb{E}_{\mathbb{P}_t}\left[(\mathbf{p}_t - \boldsymbol{\mu}_t)(\mathbf{p}_t - \boldsymbol{\mu}_t)^T\right] = \boldsymbol{\Sigma}_t$ are the mean and covariance matrix of the random vector of $\mathbf{p}_t$, respectively, for ease of notation. Now we only need to transform the constraints (3.2) into equivalent deterministic formulations.

Lemma 3.1. *If the estimates $\bar{\boldsymbol{\mu}}_t$ and $\bar{\boldsymbol{\Sigma}}_t$ are exactly known, i.e., $\boldsymbol{\mu}_t = \bar{\boldsymbol{\mu}}_t$, $\boldsymbol{\Sigma}_t = \bar{\boldsymbol{\Sigma}}_t$, and $\mathcal{P}_t = \mathcal{P}_t^1$, then for $t = 1, \cdots, T$, the inequalities (3.6) and (3.7) are equivalent to the following inequalities:*

$$\begin{aligned}
& \sqrt{\frac{\varepsilon}{1-\varepsilon}} \left(\sum_{i=1}^{t-1} (T_i^t)^2 \boldsymbol{u}_{i-1}^T \bar{\boldsymbol{\Sigma}}_{i-1} \boldsymbol{u}_{i-1} + \boldsymbol{u}_{t-1}^T \bar{\boldsymbol{\Sigma}}_{t-1} \boldsymbol{u}_{t-1} \right)^{\frac{1}{2}} \\
& + \left(\underline{x} - \left(T_0^t x_0 + \sum_{i=1}^{t-1} T_i^t \boldsymbol{\mu}_{i-1}^T \boldsymbol{u}_{i-1} + \boldsymbol{\mu}_{t-1}^T \boldsymbol{u}_{t-1} \right) \right) \leq 0. \qquad (3.10)
\end{aligned}$$

$$\sqrt{\frac{\varepsilon}{1-\varepsilon}}\left(\sum_{i=1}^{t-1}(T_i^t)^2\boldsymbol{u}_{i-1}^T\bar{\boldsymbol{\Sigma}}_{i-1}\boldsymbol{u}_{i-1}+\boldsymbol{u}_{t-1}^T\bar{\boldsymbol{\Sigma}}_{t-1}\boldsymbol{u}_{t-1}\right)^{\frac{1}{2}}$$
$$+\left(\sum_{i=1}^{n}u_t^i-\left(T_0^tx_0+\sum_{i=1}^{t-1}T_i^t\boldsymbol{\mu}_{i-1}^T\boldsymbol{u}_{i-1}+\boldsymbol{\mu}_{t-1}^T\boldsymbol{u}_{t-1}\right)\right)\leq 0. \quad (3.11)$$

Proof: We can rewrite the inequalities (3.6) as

$$\min_{\mathbb{P}_t\in\mathcal{P}_t}\mathbb{P}_t(\underline{x}-x_t\leq 0)\geq 1-\varepsilon.$$

In light of Theorem 2.2 in [103], we know that

$$\min_{\mathbb{P}_t\in\mathcal{P}_t}\mathbb{P}_t(\underline{x}-x_t\leq 0)\geq 1-\varepsilon\Longleftrightarrow\sup_{\mathbb{P}_t\in\mathcal{P}_t}\mathbb{P}_t-\mathrm{CVaR}_\varepsilon(\underline{x}-x_t)\leq 0.$$

By virtue of Lemma 2.2 in [10], we have

$$\begin{aligned}
&\sup_{\mathbb{P}_t\in\mathcal{P}_t}\mathbb{P}_t-\mathrm{CVaR}_\varepsilon(\underline{x}-x_t)\\
=&\min_{\alpha}\sup_{\zeta\sim(\underline{x}-\mathbb{E}(x_t),\mathrm{Var}(x_t))}\alpha+\frac{1}{1-\epsilon}\mathbb{E}\left[(-\alpha-\zeta)_+\right]\\
=&\min_{\alpha}\alpha+\frac{1}{1-\varepsilon}\sup_{\zeta\sim(\underline{x}-\mathbb{E}(x_t),\mathrm{Var}(x_t))}\mathbb{E}\left[(-\alpha-\zeta)_+\right]\\
=&\min_{\alpha}\left\{\alpha+\frac{1}{2(1-\varepsilon)}\left[\sqrt{\mathrm{Var}(x_t)+(\underline{x}-\mathbb{E}(x_t)+\alpha)^2}-(\alpha+\underline{x}-\mathbb{E}(x_t))\right]\right\}.
\end{aligned}$$

Define

$$h_\varepsilon(\alpha)=\alpha+\frac{1}{2(1-\varepsilon)}\left[\sqrt{\mathrm{Var}(x_t)+(\underline{x}-\mathbb{E}(x_t)+\alpha)^2}-(\alpha+\underline{x}-\mathbb{E}(x_t))\right].$$

Let $\frac{\partial h_\varepsilon(\alpha)}{\partial\alpha}=0$, we obtain

$$\alpha^*=\frac{2\varepsilon-1}{2\sqrt{\varepsilon(1-\varepsilon)}}\sqrt{\mathrm{Var}(x_t)}+\underline{x}-\mathbb{E}(x_t).$$

Substituting α^* into $h_\varepsilon(\alpha)$, we obtain

$$\begin{aligned}
&\sup_{\mathbb{P}_t\in\mathcal{P}_t}\mathbb{P}_t-\mathrm{CVaR}_\varepsilon(\underline{x}-x_t)\\
=&\sqrt{\frac{\varepsilon}{1-\varepsilon}}\left(\sum_{i=1}^{t-1}(T_i^t)^2\mathbf{u}_{i-1}^T\bar{\boldsymbol{\Sigma}}_{i-1}\mathbf{u}_{i-1}+\mathbf{u}_{t-1}^T\bar{\boldsymbol{\Sigma}}_{t-1}\mathbf{u}_{t-1}\right)^{\frac{1}{2}}\\
&-\left(\underline{x}-\left(T_0^tx_0+\sum_{i=1}^{t-1}T_i^t\boldsymbol{\mu}_{i-1}^T\mathbf{u}_{i-1}+\boldsymbol{\mu}_{t-1}^T\mathbf{u}_{t-1}\right)\right)
\end{aligned}$$

Similarly, we can prove the inequalities (3.11). The proof is completed. ■

Combining Lemma 3.1 and the equality (3.8) yields the following theorem:

Theorem 3.1. *Problem DRMPS with $\mathcal{P}_t = \mathcal{P}_t^1$ is equivalent to the following optimization problem:*

$$\max \quad T_0^T x_0 + \sum_{i=1}^{T-1} T_i^t \boldsymbol{\mu}_{i-1}^T \boldsymbol{u}_{i-1} + \boldsymbol{\mu}_{T-1}^T \boldsymbol{u}_{T-1} \tag{3.12}$$

$$\begin{aligned} s.t. \quad & \sqrt{\frac{\varepsilon}{1-\varepsilon}} \left(\sum_{i=1}^{t-1} (T_i^t)^2 \boldsymbol{u}_{i-1}^T \bar{\boldsymbol{\Sigma}}_{i-1} \boldsymbol{u}_{i-1} + \boldsymbol{u}_{t-1}^T \bar{\boldsymbol{\Sigma}}_{t-1} \boldsymbol{u}_{t-1} \right)^{\frac{1}{2}} \\ + & \left(\underline{x} - \left(T_0^t x_0 + \sum_{i=1}^{t-1} T_i^t \boldsymbol{\mu}_{i-1}^T \boldsymbol{u}_{i-1} + \boldsymbol{\mu}_{t-1}^T \boldsymbol{u}_{t-1} \right) \right) \le 0, \ t = 1, \cdots, T, \end{aligned} \tag{3.13}$$

$$\begin{aligned} & \sqrt{\frac{\varepsilon}{1-\varepsilon}} \left(\sum_{i=1}^{t-1} (T_i^t)^2 \boldsymbol{u}_{i-1}^T \bar{\boldsymbol{\Sigma}}_{i-1} \boldsymbol{u}_{i-1} + \boldsymbol{u}_{t-1}^T \bar{\boldsymbol{\Sigma}}_{t-1} \boldsymbol{u}_{t-1} \right)^{\frac{1}{2}} \\ + & \left(\sum_{i=1}^{n} u_t^i - \left(T_0^t x_0 + \sum_{i=1}^{t-1} T_i^t \boldsymbol{\mu}_{i-1}^T \boldsymbol{u}_{i-1} + \boldsymbol{\mu}_{t-1}^T \boldsymbol{u}_{t-1} \right) \right) \\ & \le 0, \ t = 1, \cdots, T-1. \end{aligned} \tag{3.14}$$

Based on Theorem 3.1, Problem DRMPS has been transformed into an equivalent SOCP problem which can be easily solved. Now we study the case $\mathcal{P} = \mathcal{P}_t^2$. In the following discussions, we further assume that all the wealth is invested in the risky market without cash keeping. Thus, the dynamics of the wealth (3.1) becomes

$$x_{t+1} = \sum_{i=1}^{n} e_t^i u_t^i = \mathbf{e}_t^T \mathbf{u}_t. \tag{3.15}$$

Under this environment, Problem DRMWC should be proposed as follows:

$$\text{DRMPS:} \quad \max_{\mathbf{u}} \inf_{\mathbb{P}_t \in \mathcal{P}_t^2} \mathbb{E}(x_T) \tag{3.16}$$

$$\text{s.t.} \quad x_{t+1} = \mathbf{e}_{t+1}^T \mathbf{u}_t$$

$$\inf_{\mathbb{P}_t \in \mathcal{P}_t^2} \mathbb{P}_t(x_t \ge \underline{x}) \ge 1 - \varepsilon, \quad t = 1, \cdots, T, \tag{3.17}$$

$$\inf_{\mathbb{P}_t \in \mathcal{P}_t^2} \mathbb{P}_t \left(\mathbf{u}_t - \mathbf{e}_t^T \mathbf{u}_{t-1} \ge 0 \right) \ge 1 - \varepsilon, \quad t = 1, \cdots, T-1. \tag{3.18}$$

Clearly, solving Problem DRMPS with $\mathcal{P}_t = \mathcal{P}_t^2$ is more difficult than $\mathcal{P}_t = \mathcal{P}_t^1$ if $\boldsymbol{\mu}_t$ and $\boldsymbol{\Sigma}_t$ cannot be accessed perfectly. To circumvent this difficulty, we need to decompose Problem DRMPS with $\mathcal{P}_t = \mathcal{P}_t^2$ as a two layer optimization problem in which the inner layer is the problem with $\mathcal{P}_t = \mathcal{P}_t^1$ while the outer layer is to handle estimation inaccuracy of $\boldsymbol{\mu}_t$ and $\boldsymbol{\Sigma}_t$. To further our discussion, we need the following lemma which is a variation of Lemma 3.3 in [98]:

Lemma 3.2. *Let s be a random vector in $\mathbb{R}^J$ and let ξ be a random variable in $\mathbb{R}$. For a given $y \in \mathbb{R}^J$, let ambiguity sets $\mathcal{D}_s$ and $\mathcal{D}_\xi$ be as follows:*

$$\mathcal{D}_s = \left\{\mathbb{P} \in \mathcal{P}(\mathbb{R}^J) : \mathbb{E}_{\mathbb{P}}(s)^T \Sigma^{-1} \mathbb{E}_{\mathbb{P}}(s) \leq \gamma_1,\ \mathbb{E}_{\mathbb{P}}(ss^T) \preceq \gamma_2 \Sigma \right\}$$

$$\mathcal{D}_\xi = \left\{\mathbb{P} \in \mathcal{P}(\mathbb{R}) : \|\mathbb{E}_{\mathbb{P}}(\xi)\| \leq \sqrt{\gamma_1}\sqrt{y^T \Sigma y},\ \mathbb{E}_{\mathbb{P}}(\xi^2) \preceq \gamma_2 y^T \Sigma y \right\}$$

Then, we have

$$\inf_{\mathcal{P} \in \mathcal{D}_s} \mathbb{E}_{\mathbb{P}}(y^T s) = \inf_{\mathcal{P} \in \mathcal{D}_\xi} \mathbb{E}_{\mathbb{P}}(\xi).$$

Furthermore,

$$\inf_{\mathcal{P} \in \mathcal{D}_s} \mathbb{E}_{\mathbb{P}}(y^s) = \begin{cases} -\sqrt{\gamma_1}\sqrt{y^T \Sigma y}, & \text{if } \gamma_1 \leq \gamma_2, \\ -\sqrt{\gamma_2}\sqrt{y^T \Sigma y}, & \text{otherwise.} \end{cases} \tag{3.19}$$

Proof: For any $\mathbb{P} \in \mathcal{D}_s$, define $\xi = y^T s$. It yields that

$$\mathbb{E}_{\mathbb{P}}(\xi) = \mathbb{E}_{\mathbb{P}}(y^T s) = y^T \mathbb{E}_{\mathbb{P}}(s) \leq \|\Sigma^{\frac{1}{2}} y\| \|\Sigma^{-\frac{1}{2}} \mathbb{E}_{\mathbb{P}}(s)\|. \tag{3.20}$$

Since $s \in \mathcal{D}_s$, $\|\Sigma^{-\frac{1}{2}} \mathbb{E}_{\mathbb{P}}(s)\| \leq \sqrt{\gamma_1}$. Replacing this inequality into (3.20) yields that

$$\mathbb{E}_{\mathbb{P}}(\xi) \leq \sqrt{\gamma_1}\|\Sigma^{\frac{1}{2}} y\| = \sqrt{\gamma_1}\sqrt{y^T \Sigma y}.$$

In a similar way, we can prove the following inequality:

$$\mathbb{E}_{\mathbb{P}}(\xi) \geq -\sqrt{\gamma_1}\|\Sigma^{\frac{1}{2}} y\| = -\sqrt{\gamma_1}\sqrt{y^T \Sigma y}.$$

Thus, for any $s \in \mathcal{D}_s$ we have $\xi = y^T s \in \mathcal{D}_\xi$. It implies that

$$\inf_{\mathcal{P} \in \mathcal{D}_s} \mathbb{E}_{\mathbb{P}}(y^T s) \leq \inf_{\mathcal{P} \in \mathcal{D}_\xi} \mathbb{E}_{\mathbb{P}}(\xi).$$

On the contrary, for any $\xi \in \mathcal{D}_\xi$, define $s = \xi \Sigma y / y^T \Sigma y$, then

$$\mathbb{E}_{\mathbb{P}}(s)^T \Sigma^{-1} \mathbb{E}_{\mathbb{P}}(s) = (\mathbb{E}_{\mathbb{P}}(\xi))^2 y^T \Sigma \Sigma^{-1} \Sigma y / (y^T \Sigma y)^2 = \frac{(\mathbb{E}_{\mathbb{P}}(\xi))^2}{y^T \Sigma y} \leq \gamma_1.$$

Meanwhile,

$$\mathbb{E}_{\mathbb{P}}(ss^T) = \mathbb{E}_{\mathbb{P}}\left\{\xi^2 \frac{\Sigma y}{y^T \Sigma y} \frac{y^T \Sigma}{y^T \Sigma y}\right\} = \mathbb{E}_{\mathbb{P}}(\xi^2)\left\{\frac{\Sigma y}{y^T \Sigma y} \frac{(\Sigma y)^T}{y^T \Sigma y}\right\} \leq \gamma_2 y^T \Sigma y \frac{\Sigma y (\Sigma y)^T}{(y^T \Sigma y)^2}.$$

Furthermore, we claim that $\Sigma y (\Sigma y)^T \leq y^T \Sigma y \Sigma$, we have $\mathbb{E}_{\mathbb{P}}(ss^T) \leq \gamma_2 \Sigma$. Indeed, for any

$z \in \mathbb{R}^J$, we have

$$z^T \Sigma y (\Sigma y)^T z \;=\; \left[(\Sigma^{\frac{1}{2}} z)^T (\Sigma^{\frac{1}{2}} y) \right]^2 \leq \|(\Sigma^{\frac{1}{2}} z)\|^2 \|(\Sigma^{\frac{1}{2}} y)\|^2 = z^T \Sigma z y^T \Sigma y = z^T (y^T \Sigma y \Sigma) z.$$

Thus, the following inequality holds:

$$\mathbb{E}_{\mathbb{P}}(\xi^2) = y^T \mathbb{E}_{\mathbb{P}}(ss^T) y \leq \gamma_2 y^T \Sigma y.$$

Therefore, $s \in \mathcal{D}_s$ and

$$\inf_{\mathcal{P} \in \mathcal{D}_s} \mathbb{E}_{\mathbb{P}}(y^T s) \geq \inf_{\mathcal{P} \in \mathcal{D}_\xi} \mathbb{E}_{\mathbb{P}}(\xi).$$

Combining the above results, we obtain $\inf_{\mathcal{P} \in \mathcal{D}_s} \mathbb{E}_{\mathbb{P}}(y^T s) = \inf_{\mathcal{P} \in \mathcal{D}_\xi} \mathbb{E}_{\mathbb{P}}(\xi)$. In light of $\mathbb{E}_{\mathbb{P}}(\xi^2) \leq \gamma_2 y^T \Sigma y$ and $(\mathbb{E}_{\mathbb{P}}(\xi))^2 \leq (\mathbb{E}_{\mathbb{P}}(\xi^2))$, we have

$$\mathbb{E}_{\mathbb{P}}(\xi) \geq -\sqrt{\gamma_2 y^T \Sigma y}.$$

The inequality can also be attained. Therefore,

$$\inf_{\mathcal{P} \in \mathcal{D}_s} \mathbb{E}_{\mathbb{P}}(y^s) = \begin{cases} -\sqrt{\gamma_1}\sqrt{y^T \Sigma_i y}, & \text{if } \gamma_1 \leq \gamma_2, \\ -\sqrt{\gamma_2}\sqrt{y^T \Sigma_i y}, & \text{otherwise.} \end{cases} \tag{3.21}$$

We complete the proof. ■

To proceed it further, we cite Theorem 3.2 in [98] as the following lemma (Lemma 3.3) which will be used later.

Lemma 3.3. *Suppose that*

$$\mathcal{D} = \left\{ \mathbb{P} \in \mathcal{P}(\mathbb{R}^J) : \mathbb{E}_{\mathbb{P}}(s)^T \Sigma^{-1} \mathbb{E}_{\mathbb{P}}(s) \leq \gamma_1,\ \mathbb{E}_{\mathbb{P}}(ss^T) \preceq \gamma_2 \Sigma \right\} \tag{3.22}$$

Then, $\inf_{\mathbb{P} \in \mathcal{D}} \mathbb{P}\left\{ \boldsymbol{t}^T \boldsymbol{y} \leq M \right\} \geq 1 - \alpha$ *is equivalent to*

$$\bar{\boldsymbol{\mu}}^T \boldsymbol{y} + \left(\sqrt{\gamma_1} + \sqrt{\frac{1-\alpha}{\alpha}(\gamma_2 - \gamma_1)} \right) \sqrt{\boldsymbol{y}^T \Sigma \boldsymbol{y}} \leq M$$

if $\gamma_1/\gamma_2 \leq \alpha$, *and is equivalent to*

$$\bar{\boldsymbol{\mu}}^T \boldsymbol{y} + \sqrt{\frac{\gamma_2}{\alpha}} \sqrt{\boldsymbol{y}^T \Sigma \boldsymbol{y}} \leq T$$

if $\gamma_1/\gamma_2 > \alpha$.

Now, we can show that the distributional robust multi-period portfolio selection problem with uncertain moments $\mathcal{P}_t = \mathcal{P}_t^2$ can be transformed into a SOCP. From this point, Problem DRMPS with $\mathcal{P}_t = \mathcal{P}_t^2$ has the same computational complexity as that of $\mathcal{P}_t = \mathcal{P}_t^1$. It means that the uncertainty of the moments does not increase the complexity of the problem.

Theorem 3.2. *If $\gamma_1 \le \gamma_2$, Problem DRMPS with $\mathcal{P}_t = \mathcal{P}_t^2$ is equivalent to the following SOCP:*

$$\max_{\boldsymbol{u}_i} \quad \bar{\boldsymbol{\mu}}_T^T \boldsymbol{u}_{T-1} - \sqrt{\gamma_1 \boldsymbol{u}_{T-1}^T \bar{\Sigma}_T \boldsymbol{u}_{T-1}} \tag{3.23}$$

$$s.t. \quad \left(\sqrt{\gamma_1} + \sqrt{\frac{1-\varepsilon}{\varepsilon}(\gamma_2 - \gamma_1)}\right) \sqrt{\boldsymbol{u}_t^T \bar{\Sigma}_t \boldsymbol{u}_t} \le \bar{\boldsymbol{\mu}}_t^T \boldsymbol{u}_{t-1} - \underline{x},\ t = 1, \cdots, T; \tag{3.24}$$

$$\bar{\boldsymbol{\mu}}_{t-1}^T \boldsymbol{u}_{t-1} + \left(\sqrt{\gamma_1} + \sqrt{\frac{1-\varepsilon}{\varepsilon}(\gamma_2 - \gamma_1)}\right) \sqrt{\boldsymbol{u}_{t-1}^T \bar{\Sigma}_{t-1} \boldsymbol{u}_{t-1}}$$

$$\le \sum_{i=1}^n u_t^i;\ t = 2, \cdots, T-1. \tag{3.25}$$

If $\gamma_2 \le \gamma_1 \le \varepsilon\gamma_2$, Problem DRMPS with $\mathcal{P}_t = \mathcal{P}_t^2$ is equivalent to the following SOCP:

$$\max_{\boldsymbol{u}_i, \tau_i} \quad \bar{\boldsymbol{\mu}}_T^T \boldsymbol{u}_{T-1} - \sqrt{\gamma_2 \boldsymbol{u}_{T-1}^T \bar{\Sigma}_T \boldsymbol{u}_{T-1}} \tag{3.26}$$

$$s.t. \quad \left(\sqrt{\gamma_1} + \sqrt{\frac{1-\varepsilon}{\varepsilon}(\gamma_2 - \gamma_1)}\right) \sqrt{\boldsymbol{u}_t^T \bar{\Sigma}_t \boldsymbol{u}_t} \le \bar{\boldsymbol{\mu}}_t^T \boldsymbol{u}_{t-1} - \underline{x},\ t = 1, \cdots, T; \tag{3.27}$$

$$\bar{\boldsymbol{\mu}}_{t-1}^T \boldsymbol{u}_{t-1} + \left(\sqrt{\gamma_1} + \sqrt{\frac{1-\varepsilon}{\varepsilon}(\gamma_2 - \gamma_1)}\right) \sqrt{\boldsymbol{u}_{t-1}^T \bar{\Sigma}_{t-1} \boldsymbol{u}_{t-1}}$$

$$\le \sum_{i=1}^n u_t^i;\ t = 2, \cdots, T-1. \tag{3.28}$$

If $\gamma_1 > \varepsilon\gamma_2$, Problem DRMPS with $\mathcal{P}_t = \mathcal{P}_t^2$ is equivalent to the following SOCP:

$$\max_{\boldsymbol{u}_i, \tau_i} \quad \bar{\boldsymbol{\mu}}_T^T \boldsymbol{u}_{T-1} - \sqrt{\gamma_2 \boldsymbol{u}_{T-1}^T \bar{\Sigma}_T \boldsymbol{u}_{T-1}} \tag{3.29}$$

$$s.t. \quad \sqrt{\frac{\gamma_2}{\varepsilon}} \sqrt{\boldsymbol{u}_t^T \bar{\Sigma}_t \boldsymbol{u}_t} \le \bar{\boldsymbol{\mu}}_t^T \boldsymbol{u}_{t-1} - \underline{x},\ t = 1, \cdots, T; \tag{3.30}$$

$$\bar{\boldsymbol{\mu}}_{t-1}^T \boldsymbol{u}_{t-1} + \sqrt{\frac{\gamma_2}{\varepsilon}} \sqrt{\boldsymbol{u}_{t-1}^T \bar{\Sigma}_{t-1} \boldsymbol{u}_{t-1}} \le \sum_{i=1}^n u_t^i;\ t = 2, \cdots, T-1. \tag{3.31}$$

Proof. Denote $\mathbf{s}_t = \mathbf{e}_t - \bar{\boldsymbol{\mu}}_t$. Then, we can verify that

$$\mathcal{D}_{s_t} = \left\{ \mathbb{P} \in \mathcal{P}(\mathbb{R}^J) : \mathbb{E}_{\mathbb{P}}(\mathbf{s}_t)^T \Sigma^{-1} \mathbb{E}_{\mathbb{P}}(\mathbf{s}_t) \le \gamma_1,\ \mathbb{E}_{\mathbb{P}}(\mathbf{s}_t {\mathbf{s}_t}^T) \preceq \gamma_2 \Sigma \right\}$$

In light of Lemma 3.2, we have

$$\inf_{\mathcal{P}\in\mathcal{D}_s} \mathbb{E}_{\mathbb{P}}(\mathbf{s}_T^T\mathbf{u}_{T-1}) = \begin{cases} -\sqrt{\gamma_1}\sqrt{\mathbf{u}_{T-1}^T\Sigma_T\mathbf{u}_{T-1}}, & \text{if } \gamma_1 \le \gamma_2, \\ -\sqrt{\gamma_2}\sqrt{\mathbf{u}_{T-1}^T\Sigma_T\mathbf{u}_{T-1}}, & \text{otherwise.} \end{cases} \tag{3.32}$$

Since $x_T = \mathbf{s}_T^T\mathbf{u}_{T-1} + \bar{\boldsymbol{\mu}}_T^T\mathbf{u}_{T-1}$,

$$\inf_{\mathcal{P}\in\mathcal{D}_s} \mathbb{E}_{\mathbb{P}}(x_T) = \begin{cases} \bar{\boldsymbol{\mu}}_T^T\mathbf{u}_{T-1} - \sqrt{\gamma_1}\sqrt{\mathbf{u}_{T-1}^T\Sigma_T\mathbf{u}_{T-1}}, & \text{if } \gamma_1 \le \gamma_2, \\ \bar{\boldsymbol{\mu}}_T^T\mathbf{u}_{T-1} - \sqrt{\gamma_2}\sqrt{\mathbf{u}_{T-1}^T\Sigma_T\mathbf{u}_{T-1}}, & \text{otherwise.} \end{cases} \tag{3.33}$$

In a similar way, we can prove that

$$\inf_{\mathbb{P}_t\in\mathcal{P}_t^2} \mathbb{P}_t(x_t \ge \underline{x}) \ge 1-\varepsilon$$

is equivalent to

$$\begin{cases} \left(\sqrt{\gamma_1} + \sqrt{\frac{1-\varepsilon}{\varepsilon}(\gamma_2-\gamma_1)}\right)\sqrt{\mathbf{u}_t^T\bar{\Sigma}_t\mathbf{u}_t} \le \bar{\boldsymbol{\mu}}_t^T\mathbf{u}_{t-1} - \underline{x}, & \text{if } \gamma_1/\gamma_2 \le \alpha, \\ \sqrt{\frac{\gamma_2}{\varepsilon}}\sqrt{\mathbf{u}_t^T\bar{\Sigma}_t\mathbf{u}_t} \le \bar{\boldsymbol{\mu}}_t^T\mathbf{u}_{t-1} - \underline{x}, & \text{otherwise.} \end{cases}$$

We can also prove that

$$\inf_{\mathbb{P}_t\in\mathcal{P}_t^2} \mathbb{P}_t\left(\mathbf{u}_t - \mathbf{e}_{t-1}^T\mathbf{u}_{t-1} \ge 0\right)$$

is equivalent to

$$\begin{cases} \bar{\boldsymbol{\mu}}_{t-1}^T\mathbf{u}_{t-1} + \left(\sqrt{\gamma_1} + \sqrt{\frac{1-\varepsilon}{\varepsilon}(\gamma_2-\gamma_1)}\right)\sqrt{\mathbf{u}_{t-1}^T\bar{\Sigma}_{t-1}\mathbf{u}_{t-1}} \le \sum_{i=1}^n u_t^i, & \text{if } \gamma_1/\gamma_2 \le \alpha, \\ \bar{\boldsymbol{\mu}}_{t-1}^T\mathbf{u}_{t-1} + \sqrt{\frac{\gamma_2}{\varepsilon}}\sqrt{\mathbf{u}_{t-1}^T\bar{\Sigma}_{t-1}\mathbf{u}_{t-1}} \le \sum_{i=1}^n u_t^i \le \bar{\boldsymbol{\mu}}_t^T\mathbf{u}_{t-1} - \underline{x}, & \text{otherwise.} \end{cases}$$

Then, the results (3.23) - (3.31) are easily obtained through combining the above results. We completed the proof. ■

3.4 Numerical studies

In this section, we will use some numerical experiments to illustrate our proposed method and validate its efficiency.

Example 1. Let us first consider the case without moment uncertainties. For this case, we select three portfolios from Shanghai Stock exchange. The transaction data within 60 business days is used to compute μ_t and Σ_t. Let $T = 10$ and $s_t = 1.04$. We suppose that the mean and covariance are constant during this time period. Then, the corresponding

$\mu_t = [0.122, 0.206, 0.188]^T$ and

$$\Sigma_t = \begin{bmatrix} 0.0146 & 0.0187 & 0.0145 \\ 0.0187 & 0.0854 & 0.0104 \\ 0.0145 & 0.0104 & 0.0289 \end{bmatrix}$$

Let $\underline{x} = 1.15$ and $\varepsilon = 0.05$. Then, the expected return at the end of the period is $\mathbb{E}(x_T) = 8.0758$. If we adjust $\underline{x}$ from $\underline{x} = 1.15$ to $\underline{x} = 1.196$, then the expected return $\mathbb{E}(x_T) = 5.3168$. The optimal $\mathbb{E}(x_t)$ with $\underline{x} = 1.15$ and $\underline{x} = 1.196$ are depicted in Figure 4.1. From Figure 4.1, we can observe clearly that the increase of $\underline{x}$ has a significant decrease of the expected return $\mathbb{E}(x_T)$. If we set $\underline{x} = 1.2$, then no feasible solution is found.

Fig 4.2 and Fig 4.3 show the optimal $u_2(t)$ and $u_3(t)$ with $\underline{x} = 1.15$ and $\underline{x} = 1.196$. From Fig 4.2 and Fig 4.3, we can observe that the smaller $\underline{x}$, the larger $u_2(t)$ and the smaller $u_3(t)$. The reason behind is that the second portfolio has the largest investment return, but it has the largest risk from the variance perspective. The expected investment return of the third portfolio is between the first one and the second one. The increase of $\underline{x}$ means the aversion of the risk. Thus, $u_3(t)$ will be increased with the increase of $\underline{x}$ in order to reduce the risk from the investment of the second portfolios.

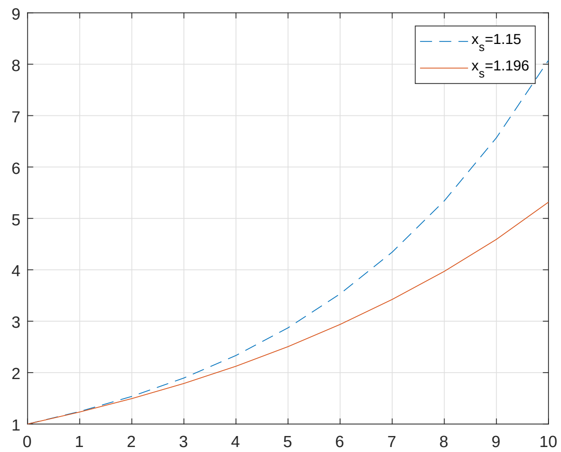


Figure 3.1: The optimal $\mathbb{E}(x_t)$ with $\underline{x} = 1.15$ and $\underline{x} = 1.196$

Example 2. Now we consider the case with uncertain moments. The expected investment return and the variance matrix are still the same as those in Example 1. The problem defined by (3.16)-(3.18) is different from the problem defined by (3.6)-(3.7) as x_t is expressed only in terms of $\mathbf{u}_t$ in the problem with uncertain moments. This expression is adopted as the problem with uncertain moments can be transferred into an equivalent second order cone programming under this formulation. We further constrain that $\sum_{i=1}^{n} u_i(t) = 1$, for

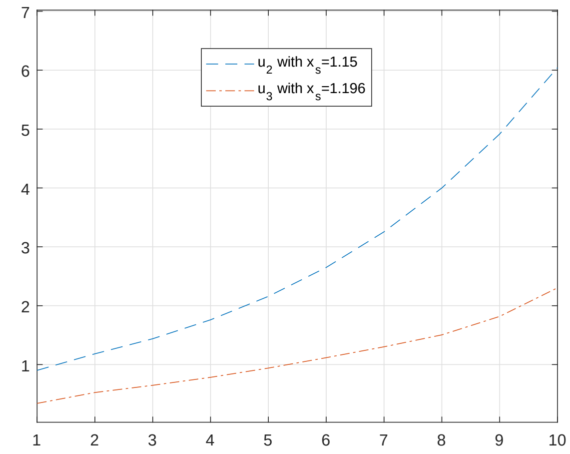


Figure 3.2: The optimal $u_2(t)$ with $\underline{x} = 1.15$ and $\underline{x} = 1.196$

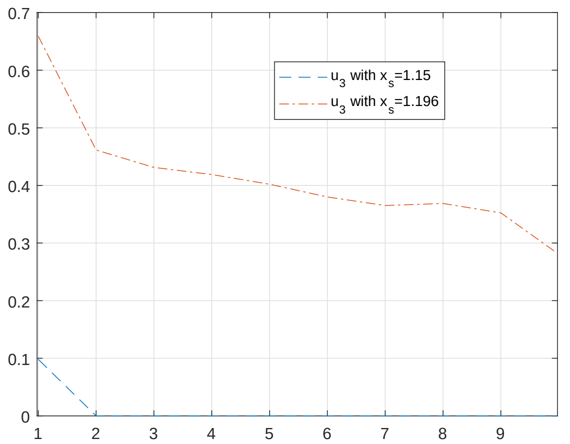


Figure 3.3: The optimal $u_3(t)$ with $\underline{x} = 1.15$ and $\underline{x} = 1.196$

all $t = 1, \cdots, T$ which will show the percentage of each portfolio invested at different times. In this case, the investment return becomes $\mu_t = [0.162, 1.246, 1.228]$. Now let $\gamma_1 = 0.0001$ and $\gamma_2 = 1.5$. At the beginning, we suppose that the portfolios are equally distributed, i.e., $u_{0,1}(0) = u_{0,2}(0) = u_{0,3}(0) = 1/3$. If we set $\underline{x} = 1$, then no feasible solution is found. Let $\underline{x} = 0.85$. The obtained $u_i(t)$, $i = 1, 2, 3$, are depicted in Fig 4.4. From the figure, we can observe that at the beginning stage, the portions of the three portfolios are similar. However, the second portfolios will increase significantly with the time evolution. The reason is that the second portfolio has the largest investment return, but it also has the largest variance. With the increase of the time, the expected worst investment returns will be large than $\underline{x}$ under the given moment uncertainty. Since the expected investment return is maximized, the portion of the second portfolio is of course becoming larger and larger.

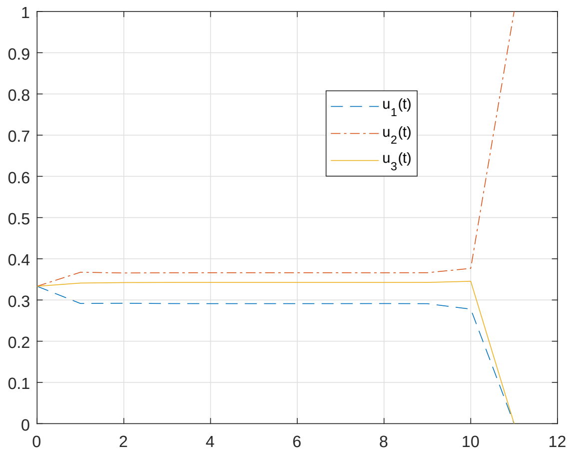


Figure 3.4: The optimal $\mathbf{u}_t$ with $\gamma_1 = 0.0001$ and $\gamma_2 = 1.2$

Now let us vary the parameters γ_1 and γ_2 to observe its impact on the optimal solution. From the definition of $\mathcal{P}_t^2$ in (3.5), we can see that the parameters γ_1 and γ_2 regulate the boundary of the uncertainties. With the increase of the γ_1 and γ_2, the uncertainty set is increased and thus, the optimal investment return under the worst case uncertainty will be decreased. Set $\gamma_1 = 0001$ and vary γ_2, the corresponding optimal investment returns under the worst distribution scenario are shown in Fig 4.5. We can clearly observe that with the increase of the parameter γ_2, the investment return is decreased.

Now we fix $\gamma_2 = 1$ and vary γ_1. From Fig 4.6, we can see that the with the increase of γ_1, the optimal investment return is also decreased. If we set $\gamma_1 = 0.01$ with $\gamma_2 = 1$, then no feasible solution is found.

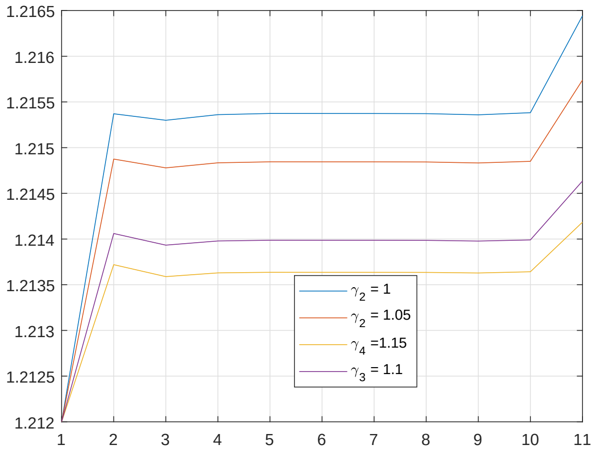


Figure 3.5: The optimal investment return under different γ_2

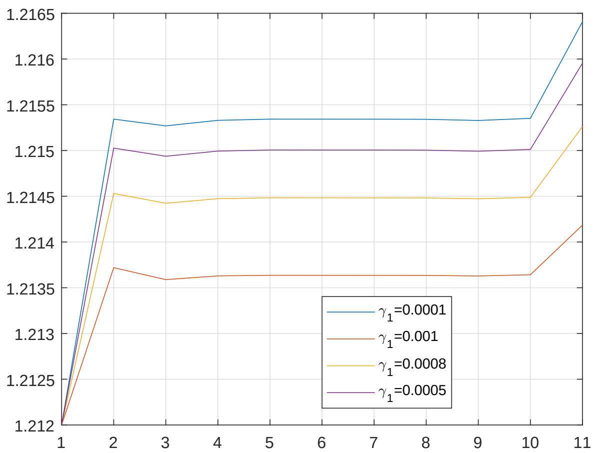


Figure 3.6: The optimal investment return under different γ_1

3.5 Conclusion

In this chapter, we have studied the dynamic portfolio selection problem with distributional uncertainty. If the moments are known exactly, then this problem can be transformed into an equivalent second order cone programming. If the moments cannot be known exactly but within a norm bounded set, we can still prove that it can be transformed into a second order cone programming. Two simple numerical examples are presented to illustrate our proposed method.

CHAPTER 4

Robust Multi-Period and Multi-Objective Portfolio Selection

4.1 Introduction

Portfolio selection is to optimally allocate investors' capital to a number of candidate securities. Traditionally, this problem has been formulated as an optimization problem through Markowitz mean-variance model [55]. In this model, the mean is used as the measurement of return investment and the variance is leveraged to measure risk. Once the expected unit return of each of the securities and their covariance matrix are given, the portfolio selection problem can be formulated as a quadratic optimization problem subject to linear constraints if the weight on the mean and the risk are given. Thus, before establishing a mean-variance model, we need to: 1) estimate the input parameters, including the expected unit return of each security and the covariance matrix and 2) determine the weights of mean and variance. In practice, the input parameters are usually estimated through empirical observations or subjective studies [34]. However, a small perturbation of the input parameters may lead to a large deviation of the selected portfolio performances. In addition, how to determine the weights of mean and variance is also challenging [34].

There are many existing works to alleviate the aforementioned problems. In [48], the minimum transaction lots is considered in Markowitz's model for portfolio selection to make the problem more practical. This problem is then formulated as a combinatorial optimization problem and a genetic algorithm is proposed to solve the problem. In [75], minimum transaction lots is further discussed with a cardinality constraint where the cardinality constraint is to constrain the number of portfolios to be selected. The problem is also formulated as a mixed-integer optimization problem and a customized genetic algorithm is introduced to solve the problem. In [16], the minimum transaction lots has been further discussed with four different models: mean-variance model, mean absolute deviation model, minimax model and combinational Value-at-Risk model. These four different models are of the form of four different discrete optimization problems. The published results show that the mean-variance model performs better than others.

How to handle uncertainties in the input parameters is also another important problem to be considered [93, 94]. In the mean-variance model, the future returns and variances are usually obtained through classic point-estimation [86]. The uncertain nature of risk and return often leads to that the results obtained are unreliable and sensitive to perturbations in parameters. There are extensive works developed to address data uncertainty in an optimization problem and this research topic is now called robust optimization. In general, there are two different types of methods to cope with uncertainty: stochastic-based methods and deterministic-based methods [6]. Stochastic-based methods usually require the statistical properties of the uncertainty. Deterministic-based method is using minimax criteria to optimize the worst case scenario and thus only the range of uncertainty is required to be known. For a quadratic programming problem with ellipsoidal uncertainties, it has been shown in [6] that it can be transformed to an equivalent conic quadratic optimization problem. If the problem is a Semi-Definite programming (SDP) and uncertainties are in an ellipsoid, the corresponding robust optimization problem is still an SDP [85]. Robust optimization is also introduced to handle the uncertainties in portfolio selection. In [69], a robust portfolio selection model with a combined worst-case conditional value-at-risk and multi-factor model is studied. The authors have shown that the probability distributions in the definition of WCVaR can be determined by specifying the mean vectors under the assumption of multivariate normal distribution with a fixed variance-covariance matrix [69]. In [31], robust mean-variance portfolio selection problem is studied. Through introducing uncertainty structures, the authors show that the robust counterpart is a second order cone program. This problem has been further discussed in [71] where two different uncertainty sets are introduced for the uncertainties of input parameters.

In the standard mean-variance model, only one period of investment is considered. However, in practice, investors are willing to adjust their investment from time to time based on real time information from a financial market. Therefore, multi-period portfolio selection has attracted much research interest. In [90], a multi-period portfolio selection problem with an uncertain investment horizon is studied. Under the assumption that the exit time follows a given distribution, the problem is transformed into one with deterministic exit time. An analytical expression of the mean-variance efficient frontier is derived. In [84], a multi-period portfolio selection problem with fixed and proportional transaction costs is considered. The optimal solution and the boundaries of the no-transaction region are obtained through introducing Lagrange multiplier and dynamic programming approach. Although extensive results are discussed in literature, there are still lack of results on the portfolio selection with transaction costs under uncertainty of input parameters. In this paper, we shall fill this gap and extend Markowitz mean-variance model to multi-period portfolio selection with transaction costs under input parameters uncertainty. We will first formulate this problem as a bi-objective optimization problem and then the robust counterpart of this problem is derived. The weighted-sum approach is introduced to present the efficient Pareto front of the formulated bi-objective optimization problem. Numerical results will be presented to demonstrate the efficiency and effectiveness of the proposed method.

4.2 Problem Statement

Consider portfolio of n financial assets which can be traded at discrete times $1, 2, \cdots, T$. Suppose that at the initial time 0, the investor has already chosen a portfolio $\mathbf{w} = (w_{1,0}, w_{w,0}, \cdots, w_{n,0})$, where $w_{i,0}$ is the investor's investment in the i-th asset for $i = 1, 2, \cdots, n$. At the beginning of each period t, $t = 1, 2, \cdots, T$, the investor needs to adjust the investment in each asset by increasing or decreasing the capital amount:

$$\Delta \mathbf{w}_t = (\Delta w_{1,t}, \Delta w_{2,t}, \cdots, \Delta w_{n,t}).$$

Let $\xi_{i,t}$ be the uncertain return rate of the i-th asset during period t. Then, at the end of period t, the investor's total wealth becomes

$$r_t = \sum_{i=1}^{n} (w_{i,t-1} + \Delta w_{i,t})(1 + \xi_{i,t}). \tag{4.1}$$

The wealth increment is thereby calculated as

$$\Delta r_t = r_t - r_{t-1}.$$

Suppose that each transaction incurs a transaction cost for asset i which is calculated as $\varsigma_i |\Delta w_{i,t}|$, where ς_i is the unit transaction cost. The total transaction cost is thereby computed as

$$\sum_{i=1}^{n} \varsigma_i |\Delta w_{i,t}|$$

Let $r_{t,f}$ be the risk-free interest rate at the time t. Then, the excess return is defined as $r_t - r_{t,f}$. Suppose that the covariance matrix of the excess returns is Σ_t. The variance $(w_{t-1} + \Delta \mathbf{w}_{t-1})^T \Sigma_t (w_{t-1} + \Delta \mathbf{w}_t)$ can be used to describe the risk of the wealth. Using these quantities, the portfolio selection at the time t can be formulated as the following optimization problem:

Problem PS:

$$\begin{aligned}
\min \quad & \left((w_{t-1}+\Delta\mathbf{w}_t)^T\Sigma_t(w_{t-1}+\Delta\mathbf{w}_t)\right)^{\frac{1}{2}}, && (4.2)\\
\max \quad & r_t-\sum_{i=1}^{n}\varsigma_i|\Delta w_{i,t}|, && (4.3)\\
\text{s.t.} \quad & w_{i,t-1}+\Delta w_{i,t}\geq 0, && (4.4)\\
& \sum_{i=1}^{n}\Delta w_{i,t}+\sum_{i=1}^{n}\varsigma_i|\Delta w_{i,t}|\leq 0. && (4.5)
\end{aligned}$$

Here the constraint (4.5) means that the investment is self-financing, i.e., the investment must be equal to or less than the net income of the sales of the assets minus the total transaction cost.

To solve Problem PS, Σ_t and $\xi_{i,t}, i=1,\cdots,n$, are required which are usually estimated from the samples of the historical data. Their estimates $\hat{\Sigma}_t$ and $\hat{\xi}_{i,t}, i=1,\cdots,n$, vary depending on the samples chosen and the method to be used for calculating them. Thus, $\hat{\Sigma}_t$ and $\hat{\xi}_{i,t}, i=1,\cdots,n$, are uncertain. Let $\xi_t=[\xi_{1,t},\cdots,\xi_{n,t}]^T$. If the samples are i.i.d and satisfy $\xi_t\sim\mathcal{N}(\bar{\xi}_t,\bar{\Sigma}_t)$, then

$$\hat{\xi}_t=\frac{1}{t}\sum_{j=0}^{t-1}\xi_j\sim\mathcal{N}(\bar{\xi}_t,\frac{1}{t}\bar{\Sigma}_t),$$

$$\hat{\Sigma}_t=\frac{1}{t-1}\sum_{j=0}^{t-1}(\xi_t-\hat{\xi}_t)(\xi_t-\hat{\xi}_t)^T\sim\mathcal{W}(\frac{1}{t-1}\bar{\Sigma}_t,t-1),$$

where $\mathcal{W}(G,v)$ denotes the Wishart distribution with scale matrix G and v degrees of freedom.

In the open literature, there are several different confidence ellipsoids introduced to achieve a robust solution. A particular one is the following separated elliptical uncertainty set proposed in [31]:

$$S_\xi=\left\{\xi_{i,t}=\xi_{i,t}^0+\varepsilon_{i,t},|\varepsilon_{i,t}|\leq\gamma_i\ i=1,\cdots,n\right\},$$

$$S_\Sigma=\left\{\Sigma_t=\Sigma_t^0+\Theta_t,\|\Theta_t\|\leq\rho\right\},$$

where $||\cdot||$ denotes the Euclidean norm, ρ and γ_i's are proper bounds estimated using historical data. A robust portfolio selection problem based on this confidence elliptical sets has been proved to achieve a robust solution with the given confidence. Instead of considering the above separated uncertainty sets, we consider a joint confidence ellipsoid as follows:

$$S_\delta(\hat{\xi}_t, \hat{\Sigma}_t) = \Big\{(\xi_t, \Sigma_t) \in \mathbb{R}^n \times \mathbb{S}^{n\times n} | t(\xi_t - \hat{\xi}_t)^T \hat{\Sigma}_t^{-1}(\xi_t - \hat{\xi}_t) + \frac{t-1}{2}\|\hat{\Sigma}_t^{-1/2}(\Sigma_t - \hat{\Sigma}_t)\hat{\Sigma}_t^{-1/2}\|_{tr}^2 \le \delta^2\Big\}, \tag{4.6}$$

where $\|A\|_{tr}^2 = tr(A^T A)$ for a matrix A and δ is a parameter characterizing the desired confidence. In this paper, the joint confidence ellipsoid uncertainty set (4.6) is used and the corresponding multi-objective optimization problem can be formally stated as:

Problem RPS:

$$\min_{\Delta \mathbf{w}_t} \max_{(\xi_t, \Sigma_t)\in S_\delta(\hat{\xi}_t, \hat{\Sigma}_t)} \left\{ \left((w_{t-1} + \Delta \mathbf{w}_t)^T \Sigma_t (w_{t-1} + \Delta \mathbf{w}_t)\right)^{\frac{1}{2}}, \ \left(\sum_{i=1}^n \varsigma_i |\Delta w_{i,t}| - r_t\right) \right\} \tag{4.7}$$

$$\text{s.t.} \quad w_{i,t-1} + \Delta w_{i,t} \ge 0 \tag{4.8}$$

$$\sum_{i=1}^n \Delta w_{i,t} + \sum_{i=1}^n \varsigma_i |\Delta w_{i,t}| \le 0 \tag{4.9}$$

(Recall r_t is a function of $\Delta \mathbf{w}_t$ and ξ_t as in (4.1).) In Problem RPS, there are two objectives. One is to minimize the worst risk and the other one is to maximize the worst wealth return over the given uncertainty set.

4.3 Problem Transformation

To solve Problem RPS, we need to transform the objective function (4.7) into a more tractable form. Note that the confidence ellipsoid uncertainty set in (4.4) is for both ξ_t and Σ_t. In this joint set, $\hat{\xi}_t$ and $\hat{\Sigma}_t$ are calculated through the available samples. As in [71], we introduce a dummy variable κ and define the following two sets:

$$S_\delta(\hat{\xi}_t) = \left\{\xi_t \in \mathbb{R}^n | t(\xi_t - \hat{\xi}_t)^T \hat{\Sigma}_t^{-1}(\xi_t - \hat{\xi}_t) \le \kappa\delta^2\right\} \tag{4.10}$$

$$S_\delta(\hat{\Sigma}_t) = \left\{\Sigma_t \in \mathbb{R}^{n\times n} | \frac{t-1}{2}\|\hat{\Sigma}_t^{-1/2}(\Sigma_t - \hat{\Sigma}_t)\hat{\Sigma}_t^{-1/2}\|_{tr}^2 \le (1-\kappa)\delta^2\right\} \tag{4.11}$$

Then, we can easily verify that Problem RPS is equivalent to the following one:

$$\min_{\Delta \mathbf{w}_t} \max_{\kappa\in[0,1]} \left\{ \max_{\Sigma_t \in S_\delta(\hat{\Sigma}_t)} \left((w_{t-1} + \Delta \mathbf{w}_t)^T \Sigma_t (w_{t-1} + \Delta \mathbf{w}_t)\right)^{\frac{1}{2}}, \ \max_{\xi_t \in S_\delta(\hat{\xi}_t)} \left(\sum_{i=1}^n \varsigma_i |\Delta w_{i,t}| - r_t\right) \right\} \tag{4.12}$$

$$\text{s.t.} \quad w_{i,t-1} + \Delta w_{i,t} \ge 0 \tag{4.13}$$

$$\sum_{i=1}^n \Delta w_{i,t} + \sum_{i=1}^n \varsigma_i |\Delta w_{i,t}| \le 0 \tag{4.14}$$

To further simplify the problem, we need to find the analytical solutions for the inner maximization problem in (4.7). More specifically, the maximization problem

$$\max_{\xi_t \in S_\delta(\hat{\xi}_t)} \left(\sum_{i=1}^{n} \varsigma_i |\Delta w_{i,t}| - r_t \right)$$

in (4.7) has the solution

$$\xi_t^* = \hat{\xi}_t - \delta \sqrt{\frac{\kappa}{n} \frac{1}{(w_{t-1} + \Delta w_t)^T \hat{\Sigma}_t (w_{t-1} + \Delta w_t)}} \hat{\Sigma}_t (w_{t-1} + \Delta w_t)$$

Using the above optimal solution, we have

$$\begin{aligned}
&\max_{\xi_t \in S_\delta(\hat{\xi}_t)} \left(\sum_{i=1}^{n} \varsigma_i |\Delta w_{i,t}| - r_t \right) \\
&= \varsigma^T |\Delta w_t| - (w_{t-1} + \Delta w_t)^T 1_n - (w_{t-1} + \Delta w_t)^T \hat{\xi}_t \\
&+ \delta \sqrt{\frac{\kappa}{n} (w_{t-1} + \Delta w_t)^T \hat{\Sigma}_t (w_{t-1} + \Delta w_t)}
\end{aligned}$$

where $1_n = [1, \cdots]^T \in \mathbb{R}^n$.

Now we consider the maximization problem $\max_{\Sigma_t \in S_\delta(\hat{\Sigma}_t)} (w_{t-1} + \Delta w_t)^T \Sigma_t (w_{t-1} + \Delta w_t)$. We are able to show that

$$\begin{aligned}
\max_{\Sigma_t \in S_\delta(\hat{\Sigma}_t)} \quad & \sqrt{(w_{t-1} + \Delta w_t)^T \Sigma_t (w_{t-1} + \Delta w_t)} \\
&= \sqrt{\left(1 + \delta \sqrt{\frac{2}{n-1} (1 - \kappa)} \right) (w_{t-1} + \Delta \mathbf{w}_t)^T \hat{\Sigma}_t (w_{t-1} + \Delta \mathbf{w}_t)} \qquad (4.15)
\end{aligned}$$

To simplify the notation, we denote

$$f_1(\Delta w_t, \kappa) = \left(1 + \delta \sqrt{\frac{2}{n-1} (1 - \kappa)} \right)^{\frac{1}{2}} \sqrt{(w_{t-1} + \Delta \mathbf{w}_t)^T \hat{\Sigma}_t (w_{t-1} + \Delta \mathbf{w}_t)}$$

$$\begin{aligned}
f_2(\Delta w_t, \kappa) \quad = \quad & \varsigma^T |\Delta w_t| - (w_{t-1} + \Delta w_t)^T 1_n - (w_{t-1} + \Delta w_t)^T \hat{\xi}_t \\
& + \delta \sqrt{\frac{\kappa}{n} (w_{t-1} + \Delta w_t)^T \hat{\Sigma}_t (w_{t-1} + \Delta w_t)}.
\end{aligned}$$

Using $f_1(\Delta w_t, \kappa)$ and $f_2(\Delta w_t, \kappa)$ defined above, we rewrite Problem RPS as follows.

Problem ERPS:

$$\min_{\Delta w_t} \max_{\kappa\in[0,1]} \quad \{f_1(\Delta w_t,\kappa), f_2(\Delta w_t,\kappa)\} \tag{4.16}$$

$$\text{s.t.} \quad w_{i,t-1} + \Delta w_{i,t} \geq 0 \tag{4.17}$$

$$\sum_{i=1}^{n} \Delta w_{i,t} + \sum_{i=1}^{n} \varsigma_i |\Delta w_{i,t}| \leq 0 \tag{4.18}$$

4.4 Solution strategy

Problem ERPS is a minimax and bi-objective optimization problem. Classical methods to solve multi-objective optimization problem include weighted-sum approach [56], $\varepsilon-$constraint approach [41], evolutionary multi-objective optimization method [19], decomposition-based approach [99], etc. The weighted-sum approach is to transform a multi-objective optimization problem into a single-objective optimization problem through a given set of weights. The Pareto-front is obtained through diversified weights. $\varepsilon-$constraint approach is to transform a multi-objective optimization problem into a single-objective optimization problem through optimizing one objective and putting all the other objectives as $\varepsilon-$constraint. Evolutionary based methods are using non-dominated sorting Genetic Algorithm to sought solutions in Pareto-front. Decomposition-based methods are also using weighting to sought Pareto solutions which is different from weighted-sum approach that maximization weight is used, rather than summation of the weight. Note that the objective function in (4.7) contains two sub-objectives and maximization is involved in the inner level. Thus, the weighted-sum approach is ideal for the problem. We now proposed such a weighted-sum method below.

4.4.1 Weighted-sum approach

Traditional weighted-sum approach is to balance the conflict objectives through weighted method. Suppose that the two weights for the two objectives in Problem ERPS are λ_1 and λ_2, respectively. Then, the original two objectives are transformed into a single one as below:

$$f_\lambda(\Delta w_t,\kappa) = \lambda_1 f_1(\Delta w_t,\kappa) + \lambda_2 f_2(\Delta w_t,\kappa) \tag{4.19}$$

where λ_1 and λ_2 are the given weights. λ_2 is usually selected as $\lambda_2 = 1-\lambda_1$.

We comment that, in practice, the values of the two objectives might not be in the same magnitude, and thus one may be relative too small than the other. In this case, we need to adjust the weight so that the points in the Pareto front can be found evenly. Let $(\Delta w_t^{1,*}, \kappa^{1,*})$ and $(\Delta w_t^{2,*}, \kappa^{2,*})$ be the optimal solutions obtained for the individual optimization problem $f_1(\Delta w_t,\kappa)$ and $f_2(\Delta w_t,\kappa)$, respectively. Then, the adjusted weights

$\widetilde{\lambda}_1$ and $\widetilde{\lambda}_2$ can be defined as:

$$\widetilde{\lambda}_1 = \frac{\frac{\lambda}{f_1(\Delta w_t^{2,*},\kappa^{2,*})-f_1(\Delta w_t^{1,*},\kappa^{1,*})}}{\frac{\lambda}{f_1(\Delta w_t^{2,*},\kappa^{2,*})-f_1(\Delta w_t^{1,*},\kappa^{1,*})}+\frac{1-\lambda}{f_2(\Delta w_t^{1,*},\kappa^{1,*})-f_2(\Delta w_t^{2,*},\kappa^{2,*})}} \tag{4.20}$$

and $\widetilde{\lambda}_2 = 1 - \widetilde{\lambda}_1$.

The benefit through this manipulation is that the normalized weights can balance the two objectives to avoid one dominates the other in the optimization process.

4.4.2 Sub-optimization solution

Solving Problem ERPS is therefore transformed into solving a series of standard optimization problems. Let us study the inner optimization problem

$$\min_{\Delta w_t} \max_{\kappa\in[0,1]} f_\lambda(\Delta w_t, \kappa) \;\text{ s.t. (4.17) and (4.18)} \tag{4.21}$$

This sub-optimization problem (4.21) is still hard to solve as both minimization and maximization are involved within the objective function. To address this difficulty, we will examine the properties of the function $f_\lambda(\Delta w_t, \kappa)$ with respect to κ. In fact, for each given $\lambda \in [0, 1]$, we have the following lemma:

Lemma 4.1. *For any given $\lambda \in [0, 1]$, the function $f_\lambda(\Delta w_t, \kappa)$ is concave with respect to κ.*

Proof: For each given $\lambda_1 = \lambda$, let $\lambda_2 = 1 - \lambda$. Then, $f_\lambda(\Delta w_t, \kappa)$ can be rewritten as:

$$\begin{aligned} f_\lambda(\Delta w_t, \kappa) \;&=\; \left(\lambda\left(1+\delta\sqrt{\frac{2}{n-1}(1-\kappa)}\right)^{\frac{1}{2}} + \delta(1-\lambda)\sqrt{\frac{\kappa}{n}}\right)\sqrt{(w_{t-1}+\Delta w_t)^T\hat{\Sigma}_t(w_{t-1}+\Delta w_t)} \\ &\quad +\varsigma^T|\Delta w_t| - (w_{t-1}+\Delta w_t)^T 1_n - (w_{t-1}+\Delta w_t)^T\hat{\xi}_t \end{aligned}$$

Define

$$g(\kappa) = \lambda\left(1+\delta\sqrt{\frac{2}{n-1}(1-\kappa)}\right)^{\frac{1}{2}} + \delta(1-\lambda)\sqrt{\frac{\kappa}{n}}$$

To prove the concaveness of $f_\lambda(\Delta w_t, \kappa)$ in terms of κ, we only need to prove $g(\kappa)$ is concave in terms of κ. For any $\kappa \in (0, 1)$, we have

$$\begin{aligned} \frac{d^2 g(\kappa)}{d\kappa^2} \;&=\; -\frac{1}{2}\frac{\delta\lambda}{(n-1)^2}\left(1+\delta\left(\frac{2}{n-1}(1-\kappa)\right)^{\frac{1}{2}}\right)^{-\frac{1}{2}}\left(\frac{2}{n-1}(1-\kappa)\right)^{-\frac{3}{2}} \\ &\quad -\frac{1}{4}\frac{\lambda\delta^2}{(n-1)^2}\left(1+\delta\left(\frac{2}{n-1}(1-\kappa)\right)^{\frac{1}{2}}\right)^{-\frac{3}{2}}\left(\frac{2}{n-1}(1-\kappa)\right)^{-1} \end{aligned}$$

$$
\begin{aligned}
& -\frac{\delta(1-\lambda)}{4n^2}\left(\frac{\kappa}{n}\right)^{-\frac{3}{2}} \\
\leq\ & 0, \text{ for all } \lambda \in [0,1] \text{ and } \kappa \in (0,1).
\end{aligned}
$$

Thus, $g(\kappa)$ is concave. We compete the proof. ∎

For each given Δw_t, δ and λ, Lemma 4.1 shows that the inner maximization problem will be maximized at its equilibrium if it is within $(0,1)$. Otherwise, it will be maximized at either $\kappa = 0$ or $\kappa = 1$. Note that

$$
\begin{aligned}
\frac{dg(\kappa)}{d\kappa} &= -\frac{1}{2}\frac{\delta\lambda}{n-1}\left(1+\delta\left(\frac{2}{n-1}(1-\kappa)\right)^{\frac{1}{2}}\right)^{-\frac{1}{2}}\left(\frac{2}{n-1}(1-\kappa)\right)^{-\frac{1}{2}} \\
&\quad +\frac{1}{2n}\delta(1-\lambda)\left(\frac{\kappa}{n}\right)^{-\frac{1}{2}} = 0 \qquad (4.22)
\end{aligned}
$$

For each given λ, δ and Δw_t, to solve $\max_{\kappa\in[0,1]} f_\lambda(\Delta w_t, \kappa)$, we first need to check whether (4.22) has a solution within (0,1). If it has a solution κ^*, then this solution also solves the problem $\max_{\kappa\in[0,1]} f_\lambda(\Delta w_t, \kappa)$ as $g(\kappa)$ is concave within $[0,1]$. Otherwise, $\kappa^* = \arg\max\{f_\lambda(\Delta w_t, 0), f_\lambda(\Delta w_t, 1)\}$. Then, the sub-optimization problem (4.21) becomes

$$
\min_{\Delta w_t} f_\lambda(\Delta w_t, \kappa^*) \ \text{ s.t. (4.17) and (4.18)} \qquad (4.23)
$$

where

$$
\begin{aligned}
f_\lambda(\Delta w_t, \kappa^*) &= g(\kappa^*)\sqrt{(w_{t-1}+\Delta w_t)^T\hat{\Sigma}_t(w_{t-1}+\Delta w_t)} \\
&\quad +\varsigma^T|\Delta w_t| - (w_{t-1}+\Delta w_t)^T 1_n - (w_{t-1}+\Delta w_t)^T\hat{\xi}_t.
\end{aligned}
$$

4.5 Numerical experiments

In this section, several examples are solved by the proposed method. We will use the numerical solutions to study the impact of uncertainties of the return rate and the variance on the multi-objective mean-variance model.

4.5.1 Convex of $g(\kappa)$

We note that the Problem PS is equivalent to Problem ERPS. In Problem ERPS, in addition to the original variable Δw_t, there is one more variable κ which only appears in $g(\kappa)$. Thus, $g(\kappa)$ plays a crucial role in the numerical solution of Problem ERPS. In the previous section, we have already showed that $g(\kappa)$ is concave with the given λ and δ. We now verify this property computationally. To show this we plot $g(\kappa)$ for $\lambda = 0.1$ and $\delta = 0.1$, 0.5 and 0.9. The results are depicted in Figure 4.2-4.3. From the three figures, we clearly see that the maximization point of $g(\kappa)$ moves from the right end-point to the middle and then to the left end-point of the interval as λ increases. This phenomenon is also observed for different δ.

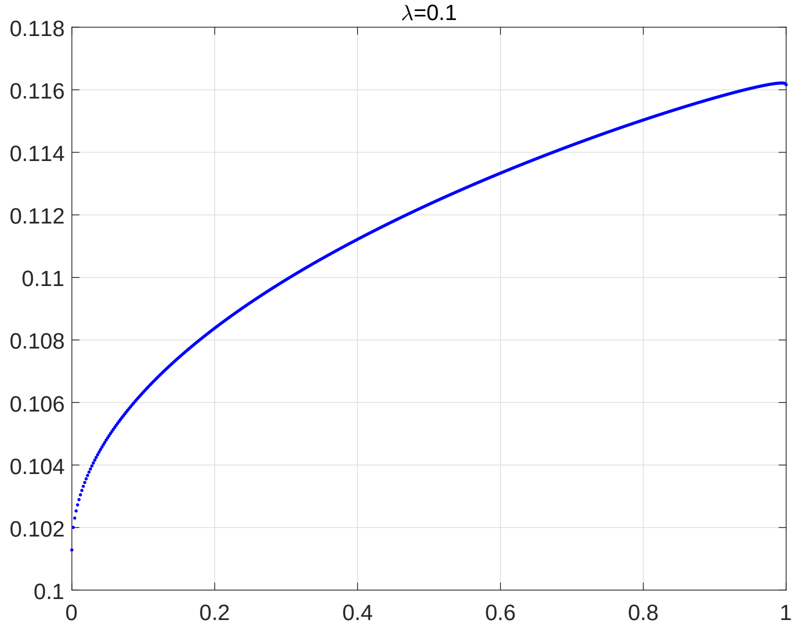


Figure 4.1: The variation of $g(\kappa)$ in terms of κ when $\lambda = 0.1$

4.5.2 Impact of the input parameters

We first study the impact of the uncertainty δ. For this purpose, we choose various values of λ and plot the optimal cost function $f_\lambda(\Delta_t^*, \kappa^*)$ against δ in Figures 4.4, 4.5, 4.6. From these figures, we see clearly that as the λ increases, the weighted optimal objective function value decreases. In fact, this phenomenon is consistent with our intuition that the larger the weight λ is, the more the investor is pursuing the return. To reduce the weight λ will lead to a more conservative investment option.

Figure 4.5 and 4.6 are shown the variation of the optimal objective function value $f_\lambda(\Delta w_t, \kappa^*)$ with respect to the uncertainty parameter δ. From the two figures, we can clearly observe that the return decreases with the increases of λ. Figure 4.5 demonstrates that $f_\lambda(\Delta w_t, \kappa^*)$ increases from -6.4598×10^4 to -6.4543×10^4 as δ goes from 0 to 0.1 when $\lambda = 0.5$, while Figure 4.6, $f_\lambda(\Delta w_t, \kappa^*)$ increases from -8.41842×10^4 to -8.417×10^4 when δ moves from 0 to 0.1 and $\lambda = 0.9$.

Figure 4.7 shows the influence of the transaction cost on the portfolio selection return. From this figure we see that, with the increase of ς from 10^{-3} to 10^{-2}, the return investment $f_\lambda(\Delta w_t, \kappa^*)$ decreases from -6.625×10^4 to -6.46×10^4. Thus, the return of the portfolio

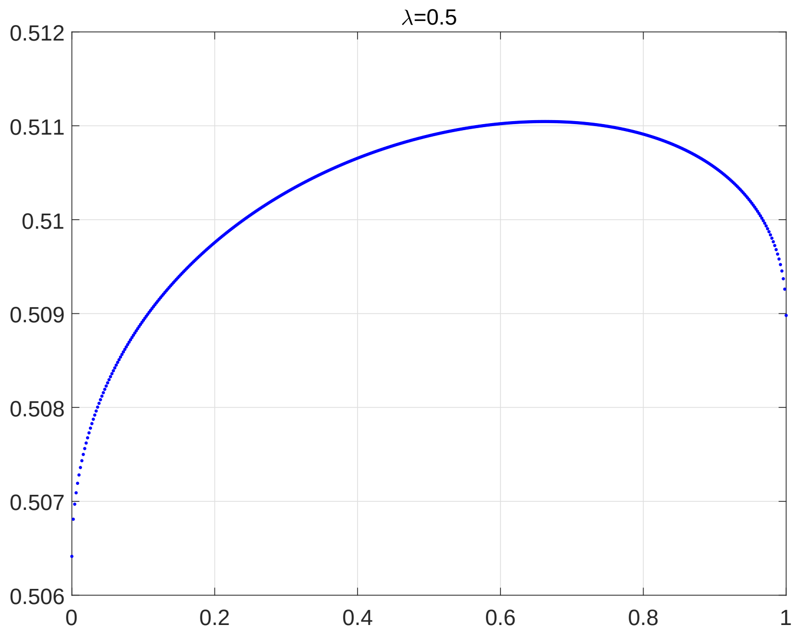


Figure 4.2: The variation of $g(\kappa)$ in terms of κ when $\lambda = 0.5$

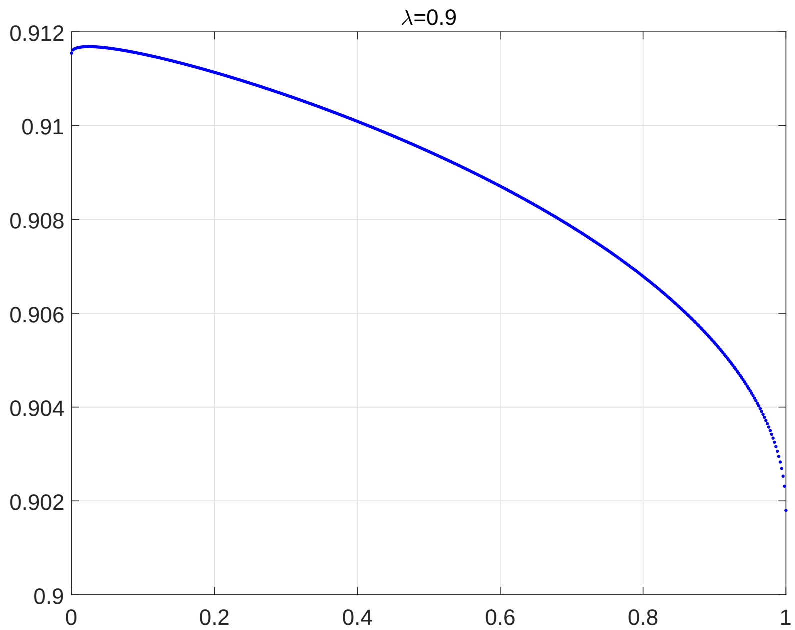


Figure 4.3: The variation of $g(\kappa)$ in terms of κ when $\lambda = 0.9$

is a decreasing function of the transaction cost ζ, as expected.

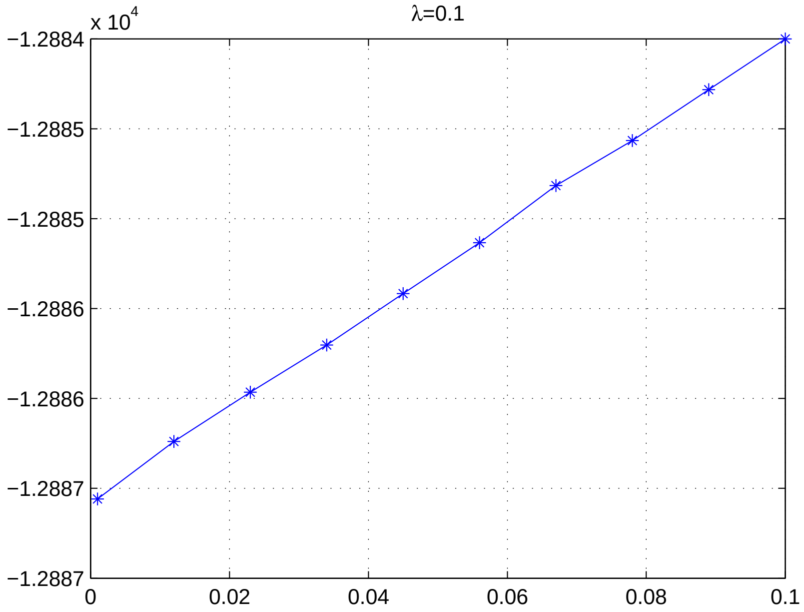


Figure 4.4: The variation of $f_\lambda(\Delta w_t, \kappa^*)$ in terms of δ when $\lambda = 0.1$

4.5.3 Numerical solutions with different parameters

In the above, we have analyzed the variation of the objective function with respect to different input parameters. We now look into the numerical solutions in different scenarios. An investment with 10 portfolios are selected in this numerical experiment. When $\delta = 0.01$ and $\varsigma = 0.01$, the obtained solutions with a 5-period investment are presented in Table 4.1. Table 4.2 and Table 4.3 show the scenarios with repsect to $(\delta, \zeta) = (0.1, 0.01)$ and $(\delta, \zeta) = (0.1, 0.001)$, respectively. From Table 4.1 and Table 4.2, we can clearly observe that transaction cost has a significant impact on the portfolio selection. For example, at the second period, selling w_1 will be changed to buying. Meanwhile, the perturbation has also impact the portfolio selection significantly. For example, for the first portfolio, it has been changed from selling to buying once we increase $\delta = 0.01$ to $\delta = 0.1$.

4.5.4 Pareto-front analysis

In this subsection, we present some results on the one-period portfolio selection under different values of the weight λ. Figure 4.8 plots the Pareto front with $\delta = 1$ and $\varsigma = 0.001$, $\varsigma = 0.002$ and $\varsigma = 0.005$, in which the horizontal axis represents $f_1(\Delta w_t, \kappa^*)$ and the vertical one represents $-f_2(\Delta w_t, \kappa^*)$. From this figure, we see clearly that $-f_2(\Delta w_t, \kappa^*)$ increases as $f_1(\Delta w_t, \kappa^*)$ decreases. From the figure we also see that when the transaction

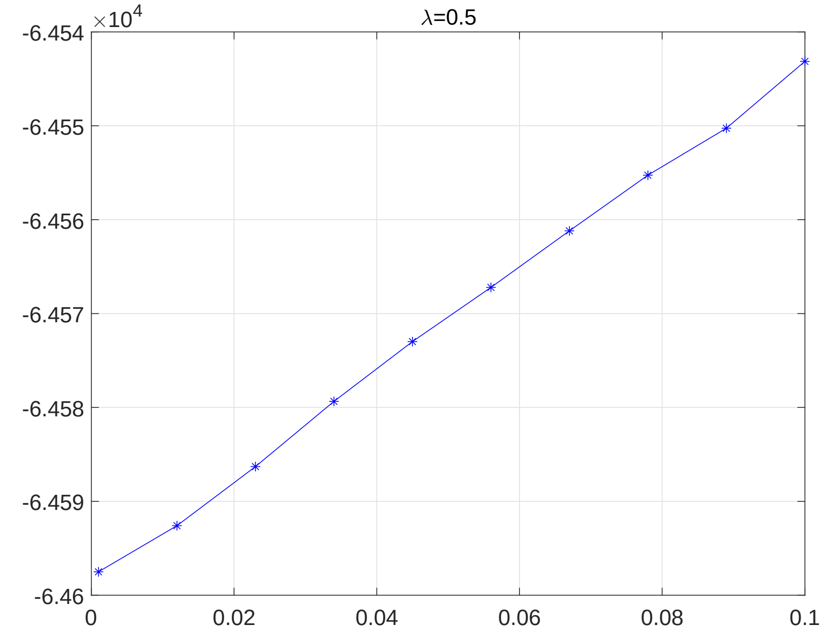


Figure 4.5: The variation of $f_\lambda(\Delta w_t, \kappa^*)$ in terms of δ when $\lambda = 0.5$

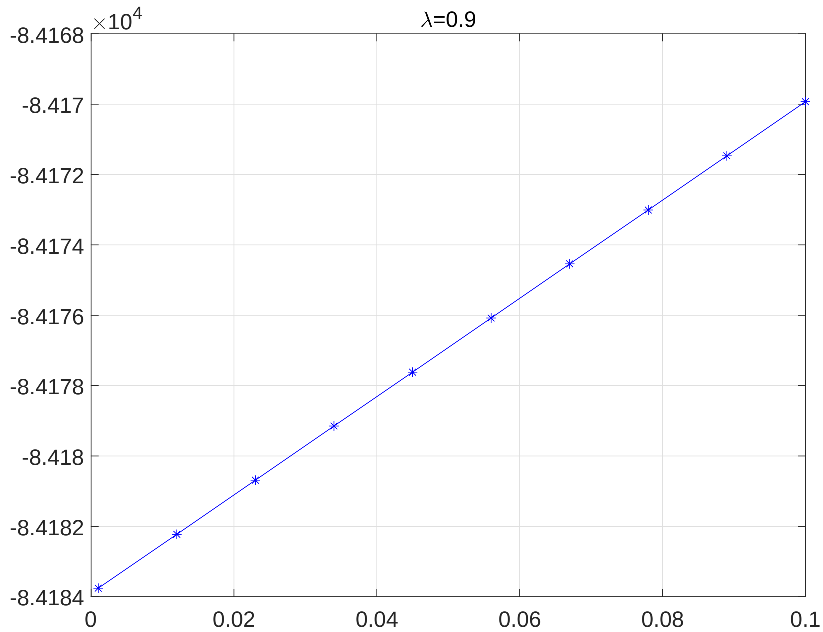


Figure 4.6: The variation of $f_\lambda(\Delta w_t, \kappa^*)$ in terms of δ when $\lambda = 0.9$

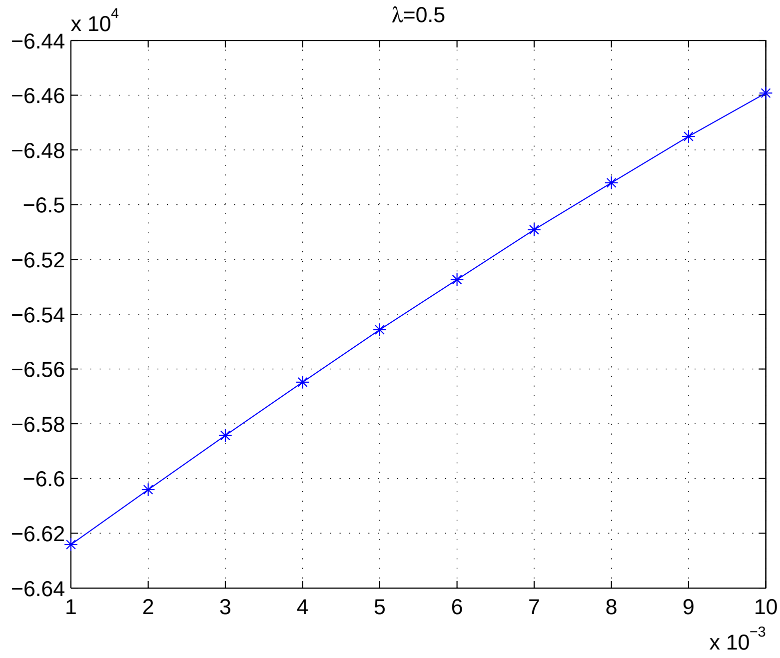


Figure 4.7: The variation of $f_\lambda(\Delta w_t, \kappa^*)$ in terms of ς when $\lambda = 0.5$

Table 4.1: Solution with five period $\varsigma = 0.001, \delta = 0.01$

	k_1	k_2	k_3	k_4	k_5
Δw_1	1.6062540e+002	1.6005381e+002	1.5941714e+002	1.5883923e+002	1.5824019e+002
Δw_2	-4.3718793e+001	-4.3458314e+001	-4.3193335e+001	-4.2953594e+001	-4.2647292e+001
Δw_3	-1.0000000e+003	-1.0000000e+003	-1.0000000e+003	-1.0000000e+003	-1.0000000e+003
Δw_4	-1.0000000e+003	-1.0000000e+003	-1.0000000e+003	-1.0000000e+003	-1.0000000e+003
Δw_5	1.9724480e+003	1.9722713e+003	1.9720766e+003	1.9718941e+003	1.9717006e+003
Δw_6	-1.3675740e+002	-1.3651295e+002	-1.3620477e+002	-1.3596164e+002	-1.3572250e+002
Δw_7	-9.7855928e+002	-9.7836228e+002	-9.7827192e+002	-9.7813905e+002	-9.7802742e+002
Δw_8	-4.6542445e+002	-4.6521487e+002	-4.6500508e+002	-4.6479322e+002	-4.6454631e+002
Δw_9	1.5230230e+003	1.5227606e+003	1.5225158e+003	1.5222595e+003	1.5219708e+003
Δw_{10}	-3.8955914e+001	-3.8854705e+001	-3.8649778e+001	-3.8458643e+001	-3.8279222e+001

Table 4.2: Solution with five period $\varsigma = 0.01, \delta = 0.01$

	k_1	k_2	k_3	k_4	k_5
Δw_1	2.1589291e-002	-3.4351365e-002	5.7805964e-002	2.3686226e-003	1.0158904e-003
Δw_2	-2.8283011e-001	-2.7323758e-004	1.6033220e-003	-3.5364786e-003	-2.6401259e-004
Δw_3	-3.3036266e+002	-3.5623656e+002	-2.4031564e+002	-3.7257428e+002	-3.4108233e+002
Δw_4	-3.0865547e+002	-2.8842569e+002	-2.9824681e+002	-2.6018254e+002	-2.7533841e+002
Δw_5	7.0391477e+002	6.9146259e+002	6.8271055e+002	6.6623463e+002	6.5626173e+002
Δw_6	-2.7364462e-002	-3.5663105e-002	-4.9939414e-002	1.0528198e-003	-3.5473078e-001
Δw_7	-1.3859574e+002	-1.1951280e+002	-2.1609022e+002	-1.0360357e+002	-1.0937146e+002
Δw_8	-1.1308908e-001	-2.4663699e-003	-2.2421419e-003	-7.7830259e-003	-1.7166429e-007
Δw_9	2.1891752e-003	1.6847602e-007	5.8640523e-002	3.0980424e-003	7.3055641e-001
Δw_{10}	-2.8203111e-007	-3.2609527e-004	-7.4650111e-008	-1.1673710e-006	-3.2439030e-003

cost increases, the investment return decreases.

Figure 4.9 plots the Pareto front for $\delta = 5$ and $\zeta = 0.001, 0.002$ and 0.005. Figure 4.9, we again see that the more risk the investors take, the more return they will benefit.

Table 4.3: Solution with five period $\varsigma = 0.01, \delta = 0.1$

	k_1	k_2	k_3	k_4	k_5
Δw_1	-5.4467604e-002	-7.0948978e-007	-2.2561076e-004	-1.0737564e-002	-1.6291721e-004
Δw_2	-1.6419802e-003	-5.6357371e-007	-7.4291123e-003	4.3365254e-003	-2.9140832e-004
Δw_3	-1.0000000e+003	-1.0000000e+003	-1.0000000e+003	-1.0000000e+003	-1.0000000e+003
Δw_4	-1.0000000e+003	-1.0000000e+003	-1.0000000e+003	-1.0000000e+003	-1.0000000e+003
Δw_5	1.7420318e+003	1.7397785e+003	1.7370980e+003	1.7349762e+003	1.7353447e+003
Δw_6	-9.7249593e-008	-3.9907242e-002	-1.8502745e-007	-4.0113434e-002	-3.8173914e-004
Δw_7	-7.9037071e+002	-7.9581203e+002	-7.9336516e+002	-7.8689932e+002	-7.9102920e+002
Δw_8	-2.1545195e+002	-2.0339581e+002	-2.0451219e+002	-2.0476930e+002	-1.9763968e+002
Δw_9	1.2043583e+003	1.2000782e+003	1.2014231e+003	1.1974970e+003	1.1941434e+003
Δw_{10}	-3.4352764e-002	-2.3689203e-007	-1.2659914e-004	-1.2481443e-007	-2.0840295e-007

Comparing Figure 4.9 with Figure 4.8, we observe that the impact of the transaction cost on the investment return becomes less significant when δ is close to 1. This trend has been further observed in Figure 4.10.

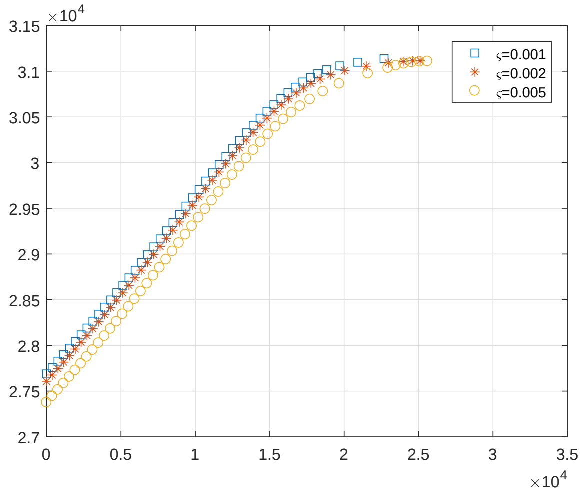


Figure 4.8: The Pareton front with $\delta = 1$

4.6 Conclusion

In this chapter, we have studied the portfolio selection problem with uncertain input parameters. Under the assumption that the uncertainty is in a ellipsoid, the original robust bi-objective optimisation problem can be transformed into an easily solved optimisation problem with the weighted sum approach which is easy to solve. The numerical results obtained demonstrate that the uncertainty of the input parameters affects the portfolio selection significantly. More specifically, the numerical results suggest that the more uncertainty in the input parameters implies that less return of the portfolio.

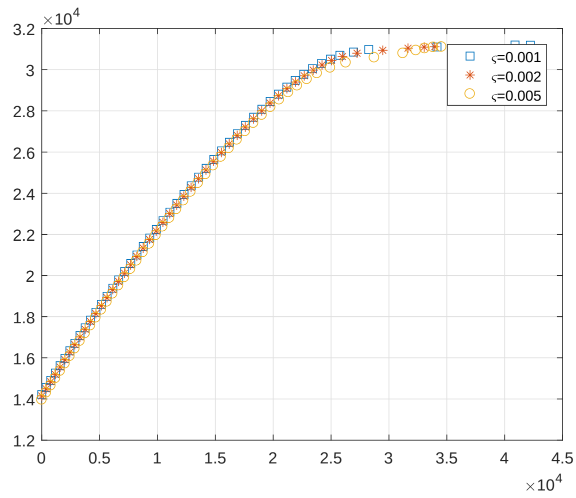


Figure 4.9: The Pareton front with $\delta = 5$

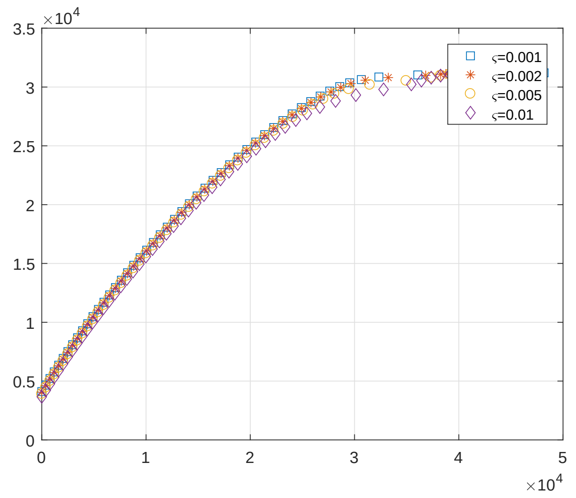


Figure 4.10: The Pareton front with $\delta = 10$

CHAPTER 5

A Nonlinear Scalarization Method for Multi-objective Optimization Problems and Applications for Portfolio Selection

5.1 Introduction

The portfolio problem has included at least two objectives: risk and return. Traditionally, the two objectives have been added together through a given weight as a combined objective function to be optimized. In practice, different investor has different preference on risk and return. The solution obtained through a weighting approach is hard to satisfy all investors. In this case, the portfolio selection problem solved by multi-objective based optimization methods is important.

Multi-objective optimization has extensive applications in engineering and management [24]. Most real-world optimization problems have multiple objectives, which can be modelled as multi-objective optimization problems (MOPs). However, due to the theoretical and computational challenges, it is not easy to solve MOPs. Therefore, multi-objective optimization has attracted a wide range of research over the last few decades.

Broadly, methods for MOPs can be categorized into three types: direct, indirect and hybrid. Population-based metaheuristic methods, such as genetic algorithm and evolutionary strategy, lend themselves to direct methods. Their iterative unit is a population instead of a single point, so they can obtain the entire set of Pareto solutions or a representative subset of it. Some typical multi-objective evolutionary algorithms (MOEAs) can be found in the works of Deb et al. [19] and Long et al. [53]. Indirect methods, mainly mean scalarization methods, reformulate MOP to a single-objective optimization problem. Normally, the Pareto solution of MOP and the optimal solution of single-objective optimization problems are corresponded. Typical indirect methods are the weighted sum method [28,33], $\varepsilon-$constraint method [43], normal-boundary intersection method [18] and Pascoletti-Serafini approach [22]. In a single run, direct methods can find an approximation of the set of Pareto solutions, while indirect methods only get one Pareto solution. Hybrid methods combine advantages of direct and indirect methods. They are based on scalar transformation and take into account heuristic ideas at the same time. A typical hybrid method is MOEA/D [96].

For indirect methods, previous research mainly focused on scalar techniques rather than on how to find an approximation of the set of Pareto solutions. This paper tries to fill this gap by extending the scalarization methods, such as the weighted sum method, to population-based. An intuitive strategy is to run a scalarization method many times using different parameters. For example, one can apply a set of different weights to the weighed sum method, and each weight will end up with a Pareto solution. However, many difficulties exist, such as how to choose weights in order to get uniformly and comprehensively distributed Pareto solutions. Furthermore, the weighted sum method works well only on convex MOPs. For nonconvex ones, nonlinear mechanisms have to be considered. Motivated by these issues, in this chapter, we are going to tackle the following topics:

(1) Extend the weighted sum method to population-based and apply it to convex MOPs;

(2) Design a population-based nonlinear scalarization method and apply it to nonconvex MOPs;

(3) Study numerical performances of the proposed linear and nonlinear scalarization methods.

(4) Solve the mean-variance-skewness portfolio selection problem by using the proposed nonlinear scalarization method.

5.2 Preliminaries

The general mathematical model of the constrained multi-objective optimization problem is as follows,

$$\text{(CMOP)} \quad \begin{cases} \text{Minimize} & \boldsymbol{F}(\boldsymbol{x}) \\ \text{Subject to} & g_i(\boldsymbol{x}) \leq 0, \quad i = 1, \cdots, p \\ & h_j(\boldsymbol{x}) = 0, \quad j = 1, \cdots, q \\ & \boldsymbol{x} \in X, \end{cases} \tag{5.1}$$

where $\boldsymbol{F} : \mathbb{R}^n \mapsto \mathbb{R}^m$ ($\boldsymbol{F}(\boldsymbol{x}) = (f_1(\boldsymbol{x}), f_2(\boldsymbol{x}), \cdots, f_m(\boldsymbol{x}))^T$) is a vector-valued function, $f_i : \mathbb{R}^n \mapsto \mathbb{R}$, $i = 1, \cdots, m$; $g_i : \mathbb{R}^n \mapsto \mathbb{R}$, $i = 1, \cdots, p$; $h_j : \mathbb{R}^n \mapsto \mathbb{R}$, $j = 1, \cdots, q$ are Lipschitz continuous functions. $X = \{\boldsymbol{x} \in \mathbb{R}^n \mid l_i \leq x_i \leq u_i\} \subset \mathbb{R}^n$ is a box set, $\boldsymbol{l} = (l_1, l_2, \cdots, l_n)^T$ and $\boldsymbol{u} = (u_1, u_2, \cdots, u_n)^T$ are lower and upper bounds, respectively.

Denote feasible set

$$\Omega = \{\boldsymbol{x} \in X \mid g_i(\boldsymbol{x}) \leq 0,\ i = 1, \cdots, p;\ h_j(\boldsymbol{x}) = 0,\ j = 1, \cdots, q\},$$

then Problem (5.1) can be simplified as

$$\begin{cases} \text{Minimize} & \boldsymbol{F}(\boldsymbol{x}) \\ \text{Subject to} & \boldsymbol{x} \in \Omega. \end{cases} \tag{5.2}$$

In multi-objective optimization, we call feasible set Ω as the *decision variable space* and its image set $\boldsymbol{F}(\Omega) = \{y = \boldsymbol{F}(\boldsymbol{x}) \mid x \in \Omega\}$ as the *objective function value space*. In the following, some definitions and theorems are reviewed.

Given two vectors

$$\boldsymbol{y} = (y_1, y_2, \cdots, y_m)^T \text{ and } \boldsymbol{z} = (z_1, z_2, \cdots, z_m)^T \in \mathbb{R}^m,$$

then

- $\boldsymbol{y} = \boldsymbol{z} \Leftrightarrow y_i = z_i$ for all $i = 1, 2, \cdots, m$;
- $\boldsymbol{y} \leq \boldsymbol{z} \Leftrightarrow y_i \leq z_i$ for all $i = 1, 2, \cdots, m$;
- $\boldsymbol{y} \prec \boldsymbol{z} \Leftrightarrow y_i < z_i$ for all $i = 1, 2, \cdots, m$;
- $\boldsymbol{y} \preceq \boldsymbol{z} \Leftrightarrow y_i \leq z_i$ for all $i = 1, 2, \cdots, m$, and $\boldsymbol{y} \neq \boldsymbol{z}$.

"$\geq$", "$\succ$" and "$\succeq$" can be defined similarly. In this paper, if $\boldsymbol{y} \preceq \boldsymbol{z}$, we say $\boldsymbol{y}$ *dominates* $\boldsymbol{z}$ or $\boldsymbol{z}$ *is dominated by* $\boldsymbol{y}$.

Definition 1. Suppose that $\boldsymbol{y} \subseteq \mathbb{R}^m$ and $\boldsymbol{y}^* \in Y$. If $\boldsymbol{y}^* \leq y$ for any $\boldsymbol{y} \in Y$, then $\boldsymbol{y}^*$ is called an *absolutely minimal point* of Y.

In the sense of minimization, absolutely minimal point is an ideal point but may not exist.

Definition 2. Let $\boldsymbol{y} \in \mathbb{R}^m$ and $\boldsymbol{y}^* \in Y$. If there is no $\boldsymbol{y} \in Y$ such that

$$\boldsymbol{y} \preceq \boldsymbol{y}^* \text{ (or } \boldsymbol{y} \prec \boldsymbol{y}^*\text{)},$$

then $\boldsymbol{y}^*$ is called an *efficient point* (or *weakly efficient point*) of Y.

The sets of absolutely minimal points, efficient points and weakly efficient points of Y are denoted as Y_{ab}, Y_{ep} and Y_{wp}, respectively. Obviously, we have $Y_{ab} \subset Y_{ep} \subset Y_{wp}$

Definition 3. Suppose that $\boldsymbol{x}^* \in \Omega$. If $\boldsymbol{F}(\boldsymbol{x}^*) \leq \boldsymbol{F}(\boldsymbol{x})$, for any $\boldsymbol{x} \in \Omega$, $\boldsymbol{x}^*$ is called an *absolutely minimal solution* of Problem (5.2). The set of absolutely minimal solution is denoted as Ω_{as}.

The concept of the absolutely minimal solution is a direct generalization of that in single-objective optimization. It is an ideal solution but may not exist for most cases.

Definition 4. Suppose that $\boldsymbol{x}^* \in \Omega$. If there is no $\boldsymbol{x} \in \Omega$ such that $\boldsymbol{F}(\boldsymbol{x}) \preceq \boldsymbol{F}(\boldsymbol{x}^*)$ (or $\boldsymbol{F}(\boldsymbol{x}) \prec \boldsymbol{F}(\boldsymbol{x}^*)$), i.e. $\boldsymbol{F}(\boldsymbol{x}^*)$ is an efficient point (or weakly efficient point) of the objective function value space $\boldsymbol{F}(\Omega)$, then $\boldsymbol{x}^*$ is called an *efficient solution* (or *weakly efficient solution*) of Problem (5.2). The sets of efficient solutions and weakly efficient solutions are denoted as Ω_{es} and Ω_{ws}, respectively.

The meaning of Pareto solution is that, if $\boldsymbol{x}^* \in \Omega_{es}$, then there is no feasible solution $\boldsymbol{x} \in \Omega$, such that any $f_i(\boldsymbol{x})$ of $\boldsymbol{F}(\boldsymbol{x})$ is not worse than that of $\boldsymbol{F}(\boldsymbol{x}^*)$ and there is at least one $i_0 \in \{1, 2, \cdots, m\}$ such that $f_{i_0}(\boldsymbol{x}) < f_{i_0}(\boldsymbol{x}^*)$. In other words, $\boldsymbol{x}^*$ is the best solution in the sense of "$\preceq$". Another intuitive interpretation of Pareto solution is that it cannot be improved with respect to any objective without worsening at least one of the others. Weakly efficient solution means that if $\boldsymbol{x}^* \in \Omega_{ws}$, then there is no feasible solution $\boldsymbol{x} \in \Omega$, such that any $f_i(\boldsymbol{x})$ of $\boldsymbol{F}(\boldsymbol{x})$ is strictly better than that of $\boldsymbol{F}(\boldsymbol{x}^*)$. In other words, $\boldsymbol{x}^*$ is the best solution in the sense of "$\prec$". The set of Pareto solutions is denoted by $\mathcal{P}^*$. Its image set $\boldsymbol{F}(\mathcal{P}^*)$ is called the *Pareto frontier*, denoted by $\mathcal{PF}^*$.

5.3 Extended weighted sum method

In this section, we consider linear scalarization methods, more specifically, the weighted sum method. Firstly, the relationship between optimal solutions of the scalarization problem and (weakly) efficient solutions of the original MOP is theoretically reviewed, then an extended weighted sum method is presented.

For Problem (5.2), consider the following scalar optimization problem

$$\text{(SOP)} \quad \begin{cases} \text{Minimize} & \sum\limits_{i=1}^{m} \lambda_i f_i(\boldsymbol{x}) \\ \text{Subject to} & \boldsymbol{x} \in \Omega, \end{cases} \tag{5.3}$$

where $\lambda = (\lambda_1, \lambda_2, \cdots, \lambda_m)^T \in \Lambda^+$ (or Λ^{++}) is a scalar vector. We call Problem (5.3) a *weighted sum scalarization* of Problem (5.2). Here

$$\Lambda^+ = \lambda = (\lambda_1, \lambda_2, \cdots, \lambda_m)^T \mid \lambda_i \geq 0 \text{ and } \sum_{i=1}^{m} \lambda_i = 1\},$$

and

$$\Lambda^{++} = \{\lambda = (\lambda_1, \lambda_2, \cdots, \lambda_m)^T \mid \lambda_i > 0 \text{ and } \sum_{i=1}^{m} \lambda_i = 1\}.$$

In the following, we theoretically analyze the relationship between Problem (5.2) and (5.3). For the sake of convenience, denote

$$\Phi_\lambda(\boldsymbol{x}) = \sum_{i=1}^{m} \lambda_i f_i(\boldsymbol{x}).$$

Definition 5. Suppose that $\Omega \in \mathbb{R}^n$ is a convex set, $\boldsymbol{F}(\boldsymbol{x}) = (f_1(\boldsymbol{x}), \cdots, f_m(\boldsymbol{x}))^T$ is a vector-valued function, if all $f_i(\boldsymbol{x})$, $i = 1, 2, \cdots, m$, are (strictly) convex on Ω, then we call $\boldsymbol{F}(\boldsymbol{x})$ an *m-dimensional (strictly) convex vector-valued function* on Ω.

Definition 6. If the feasible set Ω is convex, and the multi-objective function $\boldsymbol{F}(\boldsymbol{x})$ is a convex vector-valued function on Ω, then we call Problem (5.2) a *convex MOP*.

It is clear that $\Phi_\lambda(\boldsymbol{x})$ is convex if $\boldsymbol{F}(\boldsymbol{x})$ is convex. Therefore, Problem (5.3) is a convex problem if Problem (5.2) is a convex problem.

Theorem 5.1. For a given $\lambda \in \Lambda^{++}$ (or Λ^{+}), the optimal solution of Problem (5.3) is an efficient (or weakly efficient) solution of Problem (5.2).

Theorem 5.2. If Problem (5.2) is convex, then for any efficient solution (or weakly efficient solution) $\boldsymbol{x}^*$, there exist a $\lambda \in \Lambda^{++}$ (or $\lambda \in \Lambda^{+}$), such that $\boldsymbol{x}^*$ is an optimal solution of Problem (5.3).

Proofs of Theorems 5.1 and 5.6 can be found in the book written by Ehrgott [21]. These two theorems reveal that for a convex MOP, there is an one-to-one relationship between the weakly efficient solution of Problem (5.2) and the optimal solution of Problem (5.3). Based on this sense, we design the following extended weighted sum method.

Algorithm 1 Extended weighted sum method (EWSM)

Input: Problem parameters: f_i, g_i, h_i and X, number of solutions: N

Output: Pareto solutions: $\mathcal{P}$, Pareto frontier: $\mathcal{PF}$

// The main loops

Step 1: Generate N weights $\lambda \in \Lambda^{+}$, store them in $\bar{\Lambda}$, so $\bar{\Lambda} \subset \Lambda^{+}$.

Step 2: For each $\lambda \in \bar{\Lambda}$, globally solve Problem (5.3), the obtained optimal solution $\boldsymbol{x}^*_\lambda$ is an weakly efficient solution of Problem (5.2), store $\boldsymbol{x}^*_\lambda$ in $\mathcal{P}$.

Step 3: Compute set $F_{\bar{\Lambda}} = \{\boldsymbol{F}(\boldsymbol{x}^*_\lambda) \mid \lambda \in \bar{\Lambda}\}$, then $F_{\bar{\Lambda}}$ is an approximate Pareto frontier of Problem (5.2), let $\mathcal{PF} = F_{\bar{\Lambda}}$.

The following are some remarks about Algorithm 1:

(1) In Step 1, approaches to construct the finite subset $\bar{\Lambda}$ are various. Two intuitive approaches are presented here: (i) all $\lambda \in \Lambda^+$ consist of a simplex, so λ can be uniformly picked on this simplex; (ii) randomly pick finite number of $\lambda \in \mathbb{R}^m_+$, then normalize them to construct $\bar{\Lambda}$. More details about constructing $\bar{\Lambda}$ refers to Section 5.7.

(2) In Step 2, a global optimization method is needed to solve Problem (5.3). If Problem (5.2) is convex (so is Problem (5.3)), Problem (5.3) can be efficiently solved use any convex optimization solver. If objective functions of Problem (5.2) are nonconvex and complicated, which may lead a tough Problem (5.3), it will be not easy to solve this problem using Algorithm 1.

(3) For different $\lambda \in \bar{\Lambda}$, Problems (5.3) are independent with each other, so the parallel computing mechanism can be introduced to Algorithm 1, which will dramatically increase its efficiency. Section 5.7 will introduce more implementation details.

The geometrical explanation of the extended weighted sum method is given as follows. As shown in Figure 5.1(a), $\boldsymbol{F}(\Omega)$ is the image set of $\boldsymbol{F}(\boldsymbol{x})$ on $\Omega \subset \mathbb{R}^n$. For $\lambda \in \Lambda^+$,

$$\Phi_\lambda(\boldsymbol{x}) = \sum_{i=1}^{m} \lambda_i f_i(\boldsymbol{x}) \triangleq \sum_{i=1}^{m} \lambda_i f_i = \lambda^T f$$

is a linear function of $f = (f_1, f_2, \cdots, f_m)^T$, where $f_i = f_i(\boldsymbol{x})$, $i = 1, 2, \cdots, m$. Therefore, solving Problem (2) equals to minimize a linear function on the image set $\boldsymbol{F}(\Omega)$, i.e.,

$$\begin{cases} \text{Minimize} & \lambda^T f \\ \text{Subject to} & f \in \boldsymbol{F}(\Omega). \end{cases} \tag{5.4}$$

If f^*_λ solves Problem (5.4) for $\lambda \in \Lambda^+$, then f^*_λ must be a point in the Pareto frontier. Meanwhile, the corresponding $\boldsymbol{x}^*_\lambda$, i.e., $f^*_\lambda = \boldsymbol{F}(\boldsymbol{x}^*_\lambda)$, is an efficient solution. Although Problem (5.4) is simpler than Problem (5.3), we cannot directly work on Problem (5.4) because, first of all, the image set $\boldsymbol{F}(\Omega)$ cannot be exactly calculated; and second of all, even a solution f^*_λ is obtained, we have to solve the nonlinear equation $\boldsymbol{F}(\boldsymbol{x}) = f^*_\lambda$ to get the corresponding efficient solution $\boldsymbol{x}^*_\lambda$, which is a complex or even impossible task.

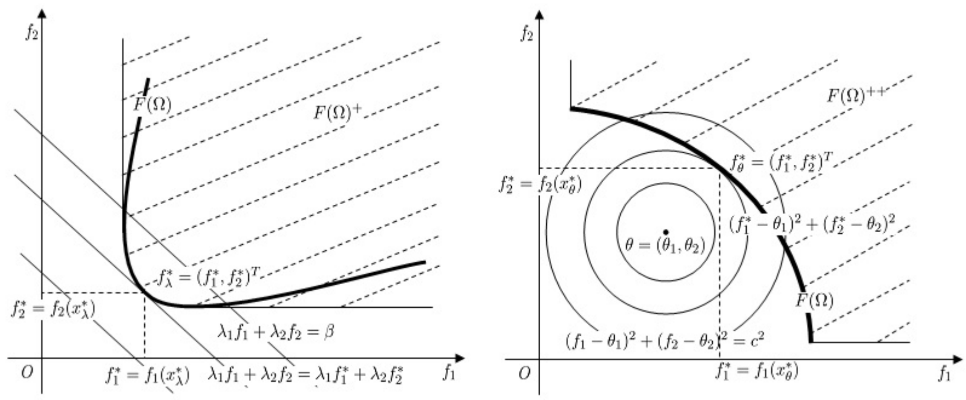
From the geometrical explanation, we can easily observe that the extended linear scalarization method works only on the problem whose image set is convex on the Pareto frontier, i.e.,

$$\boldsymbol{F}(\Omega)^+ = \{f + d \mid f \in \boldsymbol{F}(\Omega) \text{ and } d \in \mathbb{R}^m_+\}$$

is convex. Here

$$\mathbb{R}^m_+ = \{\boldsymbol{x} = (x_1, x_2, \cdots, x_m) \mid x_i \geq 0,\ i = 1, 2, \cdots, m\}.$$

If $\boldsymbol{F}(\Omega)^+$ is nonconvex (e.g., Figure 5.1(b)), only the boundary point of Pareto frontier can be obtained using Algorithm 1. However, inspired by Problem (5.4), we can nonlinearly scalarize the multi-objective function.



(a) Extended weighted sum method (b) Nonlinear scalarization method

Figure 5.1: Geometrical meaning of linear and nonlinear scalarization methods.

5.4 Nonlinear scalarization method

Nonlinear scalarization method is nothing but changing the linear objective function in Problem (5.4) into nonlinear function. A naive thought is to use a quadratic function, more specifically, m-dimensional sphere, as objective function. Therefore, we can construct the following problem

$$\begin{cases} \text{Minimize} & \sum_{i=1}^{m} (f_i - \theta_i)^2 \\ \text{Subject to} & f \in \boldsymbol{F}(\Omega), \end{cases} \tag{5.5}$$

where $f = (f_1, f_2, \cdots, f_m)^T$ and $\theta = (\theta_1, \theta_2, \cdots, \theta_m)^T \in \mathbb{R}^m$. The geometrical explanation of Problem (5.5) is shown in Figure 5.1(b). From the figure, it is possible to solve nonconvex MOPs using the nonlinear scalarization techniques. Like Problem (5.4), directly working on Problem (5.5) is out of option, but taking into account $f_i = f_i(\boldsymbol{x})$ $i = 1, 2, \cdots, m$, it can be transformed into the following problem,

$$\begin{cases} \text{Minimize} & \sum_{i=1}^{m} (f_i(\boldsymbol{x}) - \theta_i)^2 \\ \text{Subject to} & \boldsymbol{x} \in \Omega. \end{cases} \tag{5.6}$$

In the following, we discuss the relationship between the efficient solution of Problem (5.2) and the global optimal solution of Problem (5.6). For the sake of convenience, denote the objective function of Problem (5.6) as

$$\Psi_\theta(\boldsymbol{x}) = \sum_{i=1}^{m} (f_i(\boldsymbol{x}) - \theta_i)^2.$$

Suppose that $\bar{F}^* = (\bar{f}_1^*, \bar{f}_2^*, \cdots, \bar{f}_m^*)^T$, where

$$\bar{f}_i^* = \min_{x \in \Omega} \ f_i(\boldsymbol{x}), \quad i = 1, 2, \cdots, m.$$

In paper [18], $\bar{F}^*$ is called a *shadow minimum* or *utopia point*. Construct a set as

$$\bar{\Theta} = \{\theta = (\theta_1, \theta_2, \cdots, \theta_m)^T \mid \theta_i \leq \bar{f}_i^*\},$$

the element of $\bar{\Theta}$ is called a *referential point*. We can have the following theorem.

Theorem 5.3. For a given $\theta \in \bar{\Theta}$, if $\boldsymbol{x}^*$ is a global minimal solution of Problem (5.6), then $\boldsymbol{x}^*$ must be an efficient solution of Problem (5.2).

Proof: Assume that $\boldsymbol{x}^*$ is a global minimal solution of Problem (5.6) but not an efficient solution of Problem (5.2), then there exists $\bar{\boldsymbol{x}} \in \Omega$ such that $\boldsymbol{F}(\bar{\boldsymbol{x}}) \preceq \boldsymbol{F}(\boldsymbol{x}^*)$, i.e.,

$$f_i(\bar{\boldsymbol{x}}) \leq f_i(\boldsymbol{x}^*) \ \ i = 1, 2, \cdots, m; \ \text{and} \ \exists \ i_0 \ \text{s.t.} \ f_{i_0}(\bar{\boldsymbol{x}}) < f_{i_0}(\boldsymbol{x}^*).$$

So we have

$$(f_{i_0}(\bar{\boldsymbol{x}}) - \theta_{i_0})^2 < (f_{i_0}(\boldsymbol{x}^*) - \theta_{i_0})^2,$$

which yields

$$\sum_{i=1}^{m} (f_i(\bar{\boldsymbol{x}}) - \theta_i)^2 < \sum_{i=1}^{m} (f_i(\boldsymbol{x}^*) - \theta_i)^2,$$

i.e.,

$$\Psi_\theta(\bar{\boldsymbol{x}}) < \Psi_\theta(\boldsymbol{x}^*).$$

This contradicts to $\boldsymbol{x}^*$ is globally minimal, which proves the theorem.■

Remark 5.1. *Theorem 5.3 can be taken as a generalization of Theorem 5.1 for nonlinear scalarization. However, Theorem 5.6 cannot be generalized; we cannot obtain all efficient solutions through picking θ all over $\bar{\Theta}$.*

Based on Problem (5.6) and Theorem 5.3, we propose the following algorithm.

Algorithm 2 Nonlinear scalarization method (NSM)

Input: Problem parameters: f_i, g_i, h_i and X, number of solutions: N

Output: Pareto solutions: $\mathcal{P}$, Pareto frontier: $\mathcal{PF}$

// The main loops

Step 1: Successively solve

$$\bar{f}_i^* = \min_{x\in\Omega} f_i(\boldsymbol{x}), \quad i = 1, 2, \cdots, m,$$

and then construct set

$$\bar{\Theta} = \{\theta = (\theta_1, \theta_2, \cdots, \theta_m)^T \in \mathbb{R}^m \mid \theta_i \leq \bar{f}_i^*\}.$$

Step 2: Choose N referential points $\theta \in \bar{\Theta}$, store in $\hat{\Theta}$, so $\hat{\Theta} \subset \bar{\Theta}$.

Step 3: For each $\theta \in \hat{\Theta}$, globally solve Problem (5.6), the global minimal solution $\boldsymbol{x}_\theta^*$ is an efficient solution of Problem (5.2), store $\boldsymbol{x}_\theta^*$ in $\mathcal{P}$.

Step 4: Compute $F_{\hat{\Theta}} = \{\boldsymbol{F}(\boldsymbol{x}_\theta^*) \mid \theta \in \hat{\Theta}\}$, then $F_{\hat{\Theta}}$ is an approximate Pareto frontier of Problem (5.2), let $\mathcal{PF} = F_{\hat{\Theta}}$.

The following are some remarks of Algorithm 2:

(1) In Step 1, completely solving these global optimization problems is not necessary since what we really need are just lower bounds of $f_i(\boldsymbol{x})$, $i = 1, 2, \cdots, m$, so a reasonable guess of their lower bounds is enough. More practically, assume that $f_i(\boldsymbol{x}) \geq 0$, $i = 1, 2, \cdots, m$ (if any one of them is not satisfied, we can always move it parallel without changing the efficient solutions of the original problem), then we can let $\bar{\Theta} = -\mathbb{R}^m_{++}$.

(2) In Step 2, $\bar{\Theta}$ is a lowerly unbounded set, so elements of $\hat{\Theta}$ should be chosen from its upper boundary.

(3) In Step 3, a global optimization solver is needed as well as in Algorithm 1, so it could be numerically difficult if the objective function of Problem (5.6) is complicated. In this situation, the original MOP is not suitable to be solved by this algorithm.

In Step 1 of Algorithm 2, we restrict $\theta \in \bar{\Theta}$ in order to guarantee that the global minimal solution of Problem (5.6) is an efficient solution of Problem (5.1). However, based on Remark 5.1, if only choose $\theta \in \bar{\Theta}$, we may never reach some parts of the Pareto frontier. Actually, if $\theta \notin \bar{\Theta}$ but properly chosen, we can also obtain an efficient solution. This is analyzed as follows.

Construct set

$$\bar{\Theta}' = \{\theta = (\theta_1, \theta_2, \cdots, \theta_m) \in \mathbb{R}^m \mid \exists\, i_0 \in \{1, 2, \cdots, m\},\ \text{s.t. } \theta_{i_0} < \bar{f}^*_{i_0}\},$$

obviously $\bar{\Theta} \subset \bar{\Theta}'$. For a given $\theta \in \bar{\Theta}'$, assume we have $i_1 \in \{1, 2, \cdots, m\}$, such that $\theta_{i_1} < \bar{f}^*_{i_1}$, in the proof of Theorem 5.3, if we just have $i_0 = i_1$, i.e.,

$$\theta_{i_0} < \bar{f}^*_{i_0} \leq f_{i_0}(\bar{\boldsymbol{x}}) < f_{i_0}(\boldsymbol{x}^*) \tag{5.7}$$

and

$$(f_{i_0}(\bar{\boldsymbol{x}}) - \theta_{i_0})^2 - (f_{i_0}(\boldsymbol{x}^*) - \theta_{i_0})^2 < \sum_{i=1, i\neq i_0}^{m} (f_i(\boldsymbol{x}^*) - \theta_i)^2 - \sum_{i=1, i\neq i_0}^{m} (f_i(\bar{\boldsymbol{x}}) - \theta_i)^2, \tag{5.8}$$

we can still have

$$\Psi_\theta(\bar{\boldsymbol{x}}) < \Psi_\theta(\boldsymbol{x}^*),$$

which yields that $\boldsymbol{x}^*$ is an efficient solution. Of course, these conditions cannot be checked in advance, if $\theta \in \bar{\Theta}'$, but conditions (5.7) or (5.8) cannot be satisfied, the obtained global minimal solution $\boldsymbol{x}^*$ may not be an efficient solution. But we can use a non-dominated sorting [19] to exclude these points. Based on this observation, we propose the following slack nonlinear scalarization method.

Algorithm 3 Slack nonlinear scalarization method (SNSM)

Input: Problem parameters: f_i, g_i, h_i and X, number of solutions: N

Output: Pareto solutions: $\mathcal{P}$, Pareto frontier: $\mathcal{PF}$

// The main loops

Step 1: Successively solve

$$\bar{f}_i^* = \min_{x\in\Omega} f_i(\boldsymbol{x}) \quad i = 1, 2, \cdots, m,$$

and then construct set

$$\bar{\Theta}' = \{\theta = (\theta_1, \theta_2, \cdots, \theta_m) \in \mathbb{R}^m \mid \exists\ i_0 \in \{1, 2, \cdots, m\},\ \text{s.t. } \theta_{i_0} < \bar{f}_{i_0}^*\},$$

Step 2: Choose N referential points $\bar{\Theta}'$, store in $\hat{\Theta}'$, so $\hat{\Theta}' \subset \bar{\Theta}'$.

Step 3: For each $\theta \in \hat{\Theta}'$, globally solve the corresponding Problem (5.6), suppose that $\boldsymbol{x}_\theta^*$ is the global minimal solution, store in $S_{\hat{\Theta}'}$.

Step 4: Compute $F_{\hat{\Theta}'} = \{\boldsymbol{F}(\boldsymbol{x}_\theta^*) \mid \theta \in \hat{\Theta}'\}$.

Step 5: Successively check each $y \in F_{\hat{\Theta}'}$, if y is non-dominated , then store y in $\mathcal{PF}$ and its corresponding $\boldsymbol{x}_\theta^*$ in $\mathcal{P}$.

Step 5 of Algorithm 3 is actually a non-dominated ranking [19], here we pick the first Pareto frontier. The numerical comparison of Algorithm NSM and SNSM is presented in Section 6.1.

5.5 Implementation

In this section, we explain some implementation details of the proposed algorithms, including generating weights λ in Algorithm EWSM, generating referential points θ in Algorithm NSM and SNSM, and the global optimization solver for scalar optimization problems.

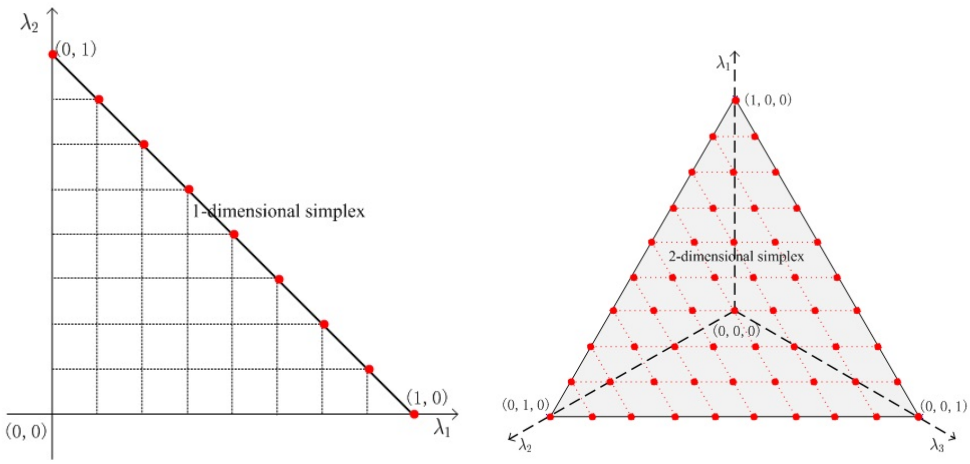
5.5.1 Generating λ in Algorithm EWSM

In Algorithm EWSM, all $\lambda \in \Lambda^+$ consist of a unit simplex, so finite number of weights should be uniformly generated in this unit simplex. The simplest strategy is to pick them randomly. For example, randomly choose $\lambda' \in [a, b]^m$, and then normalize λ' to be λ, i.e., $\lambda = \lambda' / \sum_{k=1}^m \lambda'_k$. Random strategy is simple and easy to implement, but it cannot guarantee uniformity, specially when ratio N/m is small.

In this paper, we apply a method called *systematic approach* to generate λ. This method was first introduced by Das and Dennis [18] and then applied by Deb [19]. It picks points in a normalized hyperplane (an $(m-1)$-dimensional unit simplex, which is equally inclined to all axis and has an intercept of one on each axis). If p divisions are considered along each axis, the total number of different weights N for problem with m objective functions is given by

$$N = \begin{pmatrix} m+p-1 \\ p \end{pmatrix}.$$

Detail steps about the systematic approach can be found in the work of Das and Dennis [18], here we give two examples. Figure 5.2(a) demonstrates weights when $m=2$, $p=8$, there are $N=9$ different weights; and Figure 5.2(b) demonstrates weights when $m=3$, $p=9$, there are $N=55$ different weights. It can be observe that all the weights are uniformly distributed in the unit simplex.



(a) $m=2$, $p=8$, $N=9$ (b) $m=3$, $p=9$, $N=55$

Figure 5.2: Generating weights using systematic method.

5.5.2 Generating θ in Algorithm NSM and SNSM

In Algorithm NSM, θ is generated in $\bar{\Theta}$. Figure 5.4 illustrates two strategies to generate θ when $m=2$. Here, we have

$$\boldsymbol{x}^*_{f_1} = \arg\min_{\boldsymbol{x}\in\Omega} f_1(\boldsymbol{x}), \;\; \bar{f}^*_1 = f_1(\boldsymbol{x}^*_{f_1})$$

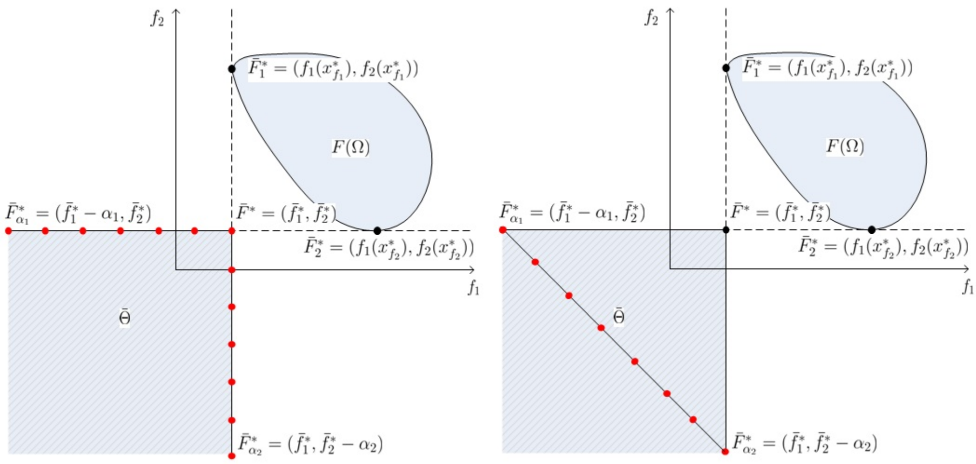
and

$$\boldsymbol{x}^*_{f_2} = \arg\min_{\boldsymbol{x}\in\Omega} f_2(\boldsymbol{x}), \;\; \bar{f}^*_2 = f_2(\boldsymbol{x}^*_{f_2}),$$

so $\bar{\Theta} = (\bar{f}_1^*, \bar{f}_2^*) - \mathbb{R}^2_{++}$. In Figure 5.3(a), referential points are uniformly generated on segment union $[(\bar{f}_1^* - \alpha_1, \bar{f}_2^*), (\bar{f}_1^*, \bar{f}_2^*)] \cup [(\bar{f}_1^*, \bar{f}_2^* - \alpha_2), (\bar{f}_1^*, \bar{f}_2^*)]$. In Figure 5.3(b), referential points are uniformly generated on segment $[(\bar{f}_1^* - \alpha_1, \bar{f}_2^*), (\bar{f}_1^*, \bar{f}_2^* - \alpha_2)]$. Here $\alpha_1, \alpha_2 > 0$ are proper positive numbers.

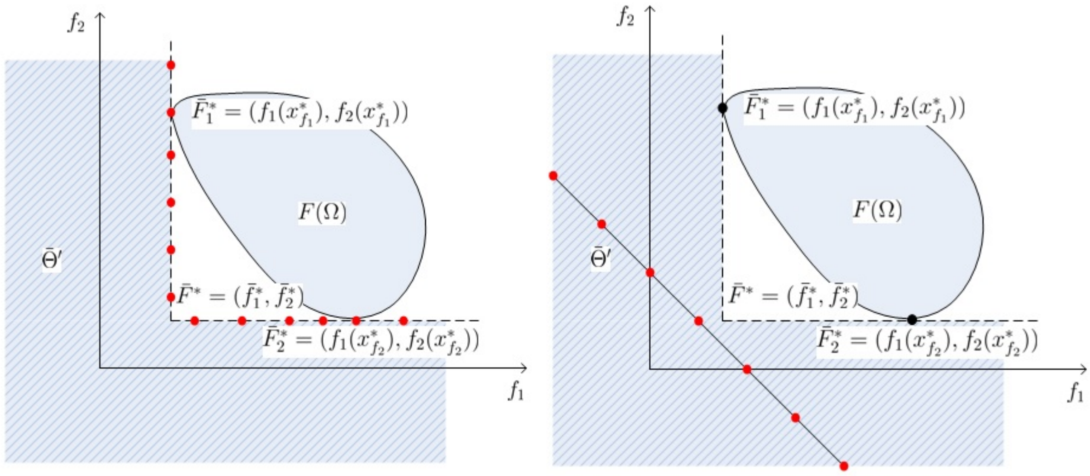
In Algorithm SNSM, θ is generated in $\bar{\Theta}'$, Figure 5.4(a) depicts the strategy of generating referential points on the upper boundary of $\bar{\Theta}'$, while Figure 5.4(b) depicts the referential points generated on a random line segment in $\bar{\Theta}'$.

When $m > 2$, the line segment becomes simplex, we can use the systematic method introduced in the previous subsection to generate referential points.



(a) θ on upper boundary (b) θ on hyperplane

Figure 5.3: Generating referential points θ for Algorithm NSM.



(a) θ on upper boundary (b) θ on hyperplane

Figure 5.4: Generating referential points θ for Algorithm SNSM.

5.5.3 Global optimization solver

The global optimization solver plays an important role in the proposed algorithms. If Problem (5.1) is convex, which leads Problems (5.3) and (5.6) convex, then they can be efficiently solved using local or global optimization solvers; otherwise, we have to use nonconvex solvers, such as the global quasisecant method and hybrid global optimization method, to tackle them. In our implementation, we simply use functions in the MATLAB optimization toolbox, such as $fmincon$, $fminsearch$ and ga, to solve Problems (5.3) and (5.6).

5.6 Mean-Variance-Skewness Portfolio Selection

The mean-variance-skewness model for portfolio selection problem can be formulated as below [91]:

$$\begin{aligned} \text{Maximize } R(x) &= x^T\bar{R} = \sum_{i=1}^{n} x_i\bar{R}_i \\ \text{Minimize } V(x) &= x^T\Sigma x = \sum_{i=1}^{n}\sum_{j=1}^{n} x_i\sigma_{ij}x_j \\ \text{Maximize } S(x) &= E(x^T(R-\bar{R}))^3 \qquad (5.9) \\ &= \sum_{i=1}^{n} x_i^3 s_i^3 + 3\sum_{i=1}^{n}(\sum_{j=1,j\neq i}^{n} x_i^2 x_j s_{iij} + \sum_{j=1,j\neq i}^{n} x_i x_j^2 s_{ijj}) \qquad (5.10) \end{aligned}$$

where $\Sigma = [\sigma_{ij}]$ is the covariance matrix and $[s_{ijk}]$ is the skewness of the candidate portfolios. Skewness is the degree of distortion from the symmetrical bell curve in a probability distribution. If skewness is negative, it shows that this portfolio has large probability for loss but small probability for profit. Suppose that $\{R_{ki}\}_{k=1}^{K}$ are the samples of the i-th portfolio, $i = 1, \cdots, n,$. Then, $\bar{R}_i$, σ_{ij} and s_{ijk} can be estimated as below:

$$\begin{aligned} \bar{R}_i &= \sum_{k=1}^{K} R_{ki}/K; \\ \sigma_{ij} &= \sum_{k=1}^{K}(R_{ki}-\bar{R}_i)(R_{kj}-\bar{R}_j)/K; \\ s_{ijk} &= \sum_{t=1}^{K}(R_{ti}-\bar{R}_i)(R_{tj}-\bar{R}_j)(R_{tk}-\bar{R}_k)/K. \end{aligned}$$

As shown in Theorem , if the third objective skewness is not involved, then we can use the linear weighting method to approximate the Pareto-front. The presence of the skewness

Table 5.1: Multi-objective test problems.

Pro.	n	Variable bounds	Objective functions	Optimal solutions	Convexity
SCH	1	$[-5, 10]$	$f_1(x) = x^2$ $f_2(x) = (x-2)^2$	$x \in [0, 2]$	convex
FON	3	$[-4, 4]$	$f_1(x) = 1 - \exp(-\sum_{i=1}^{3}(x_i - \frac{1}{\sqrt{3}})^2)$ $f_2(x) = 1 - \exp(-\sum_{i=1}^{3}(x_i + \frac{1}{\sqrt{3}})^2)$	$x_1 = x_2 = x_3$ $\in [-1/\sqrt{3}, 1/\sqrt{3}]$	nonconvex
KUR	3	$[-5, 5]$	$f_1(x) = \sum_{i=1}^{n-1}(-10\exp(-0.2\sqrt{x_i^2 + x_{i+1}^2}\ ))$ $f_2(x) = \sum_{i=1}^{n}(\|x_i\|^{0.8} + 5\sin^3(x_i))$	[19]	nonconvex

objective makes the problem become non-convex. In this chapter, we will apply our nonlinear scalarisation method to solve this multi-objective optimisation problem.

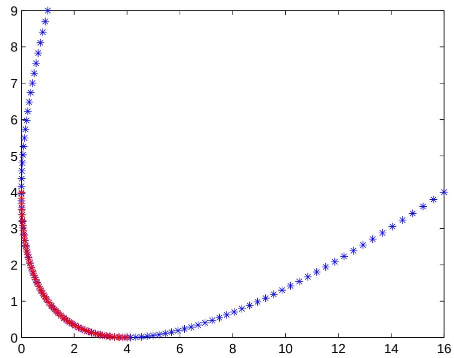
5.7 Numerical experiments

In this section, we first present some illustrative examples to demonstrate the numerical performance of the proposed algorithms, then compare the proposed algorithms with two typical heuristic multi-objective optimization solvers: NSGAII [19] and MOEA/D [96]. All the numerical experiments are implemented in an environment of MATLAB(2010a) installed on an ACER ASPIRE 4730Z laptop with a 2G RAM and a 2.16GB CPU.

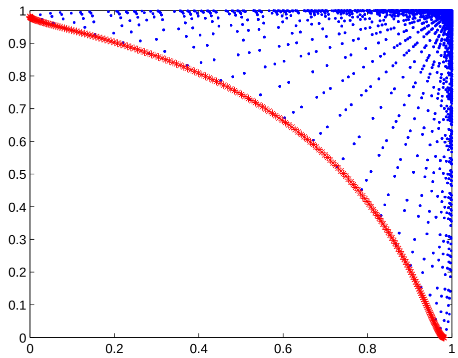
5.7.1 Illustrative examples

Problem SCH [19] in Table 5.1 is a one dimensional convex multi-objective problem. Its efficient solution set is $[0, 2]$, Figure 5.5(a) shows its image set and Pareto frontier. Solving Problem SCH using EWSM, we can obtain results showing in Figure 5.6. Among them, $\lambda \in \bar{\Lambda}$ for Figure 5.6(a) is uniformly chosen on the line segment $\lambda_1 + \lambda_2 = 1$ $(\lambda_1, \lambda_2 \geq 0)$; while $\lambda \in \bar{\Lambda}$ for Figure 5.6(b) is randomly chosen. Note that λ and Pareto points are actually in different spaces, but in Figure 5.6 (so as the following figures), we draw them together to demonstrate the relationship between λ and Pareto frontier. From Figure 5.6, we can observe that Problem SCH is perfectly solved by EWSM (the numerical performance of uniformly chosen $\bar{\Lambda}$ is better than the randomly chosen one), and each point in Pareto frontier corresponds to a $\lambda \in \bar{\Lambda}$.

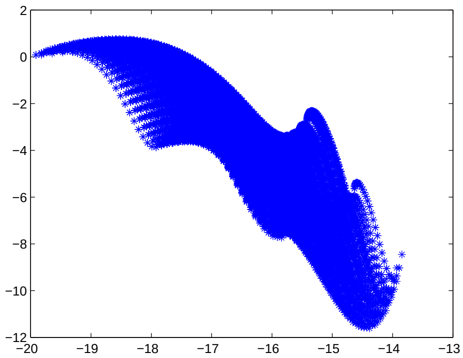
Problem FON [19] (see Table 5.1 and Figure 5.5(b)) is a three dimensional nonconvex problem, its Pareto solutions satisfy $x_1 = x_2 = x_3$, where $x_i \in [-1/\sqrt{3}, 1/\sqrt{3}]$, $i = 1, 2, 3$. Figure 5.7 demonstrates Problem FON solved by NSM. Among them, in Figure 5.7(a), referential points θ are uniformly generated on the upper boundary of $\bar{\Theta}$, i.e., $[(-1, 0), (0, 0)] \cup [(0, 0), (0, -1)]$; while in Figure 5.7(b), θ are uniformly generated on simplex $[(-1, 0), (0, -1)]$. Compare both figures, one can observe that both strategies can



(a) Problem SCH



(b) Problem FON

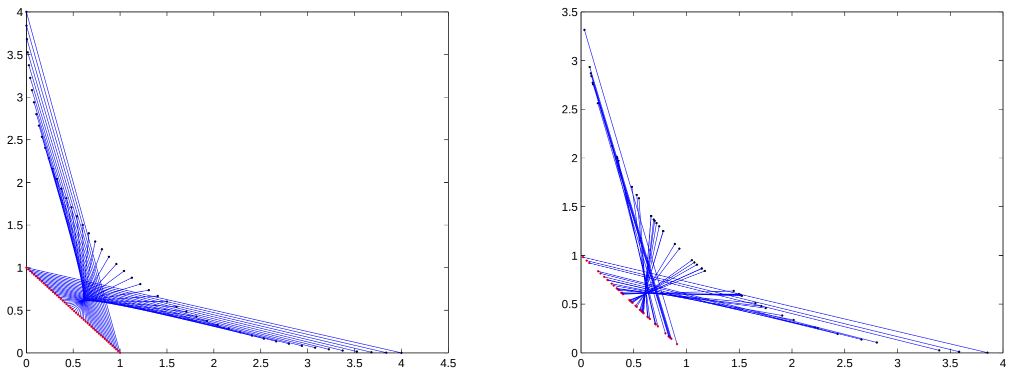


(c) Problem KUR

Figure 5.5: Objective function value set of Problems SCH, FON and KUR.

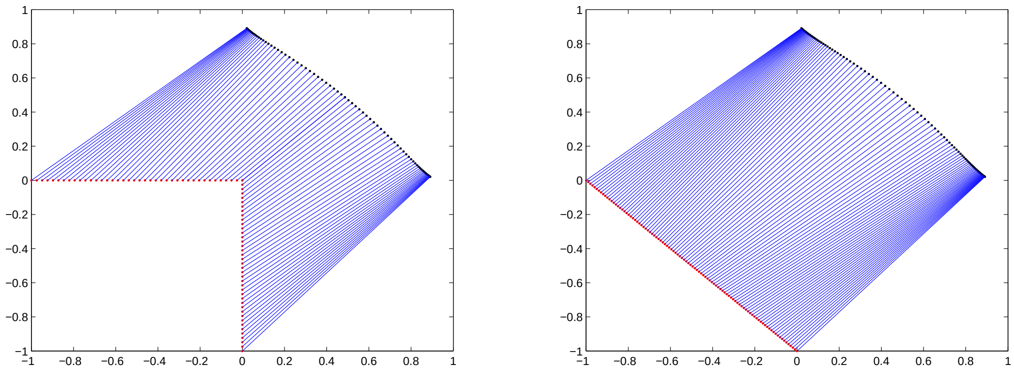
obtain perfect approximation of the Pareto frontier. Figure 5.8 demonstrates Problem FON solved by SNSM, two referential points generating strategies are applied as well. Form Figure 5.8, SNSM can still get good Pareto frontier approximation, but some non-efficient point appears at both ends of the approximate Pareto frontier. These non-efficient points can be identified and removed using Pareto sorting, which is then depicted in Figure 5.9.

Problem KUR [19] (see Table 5.1 and Figure 5.5(c)) is a three dimensional nonconvex problem, its Pareto frontier is disconnected. Figure 5.10 demonstrates Problem KUR solved by NSM and SNSM. From Figure 5.10(a), when solving by NSM, there are only a few efficient solutions can be obtained, and different $\theta \in \bar{\Theta}$ may end up with same efficient point. However, when solving using SNSM, as illustrated in Figure 5.10(b), we can see that the disconnected Pareto frontier of Problem KUR is perfectly simulated and most of the $\theta \in \bar{\Theta}'$ are properly chosen. This reveals that SNSM, although theoretically defective, could numerically performs better than NSM.



(a) $\bar{\Lambda}$ uniformly chosen. (b) $\bar{\Lambda}$ randomly chosen.

Figure 5.6: Solving Problem SCH using Algorithm 1.



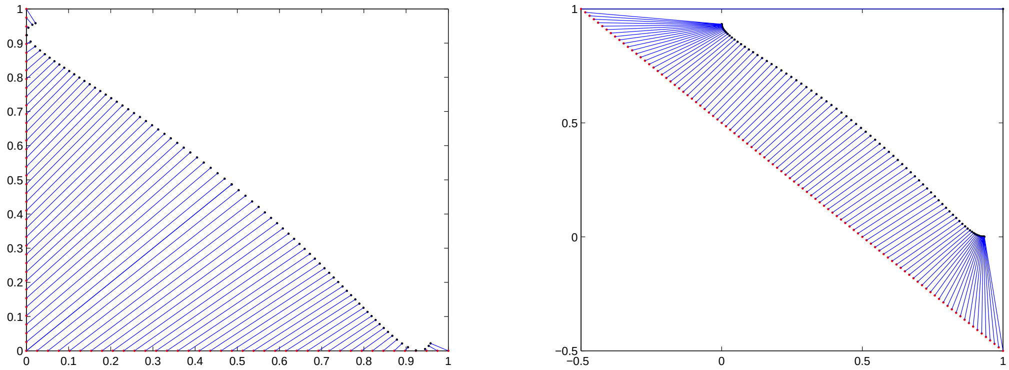
(a) λ on upper boundary (b) λ on simplex

Figure 5.7: Problem FON solved by NSM.

5.7.2 Comparison with MOEA/D and NSGAII

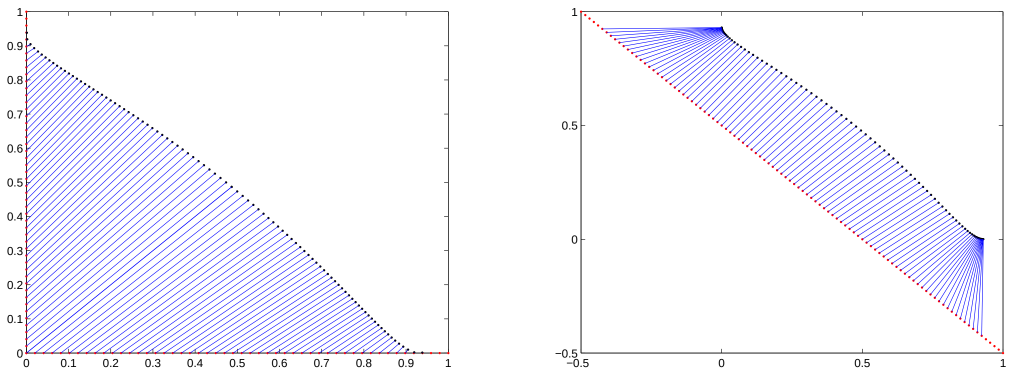
In this subsection, we first introduce MOGA/D [96] and NSGAII [19] as referential methods and then analyze the complexity and numerical performance of EWSM and SNSM by comparing them with MOEA/D and NSGAII on a set of test instances. The reason to choose MOEA/D and NSGAII as referential methods is that MOEA/D is one of the typical decomposition methods for MOPs and NSGAII is the most successful multi-objective genetic algorithm.

MOEA/D is a typical multi-objective optimization method based on evolutionary algorithm and decomposition. It decomposes an MOP into a number of scalar optimization subproblems and optimize them simultaneously. Paper [96] presented three strategies to decompose MOPs: the weighted sum approach, Tchebycheff approach and boundary intersection approach. Our proposed methods are similar with MOEA/D in decompos-



(a) λ on upper boundary (b) λ on simplex

Figure 5.8: Problem FON solved by SNSM.

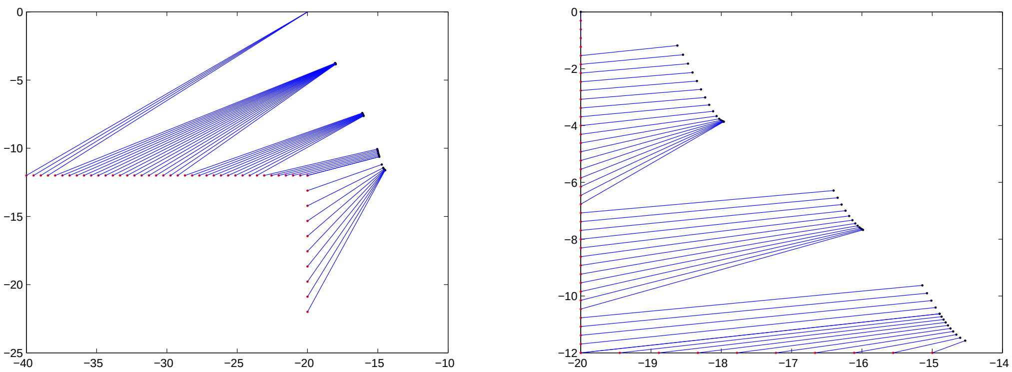


(a) λ on upper boundary (b) λ on simplex

Figure 5.9: Problem FON solved by SNSM with Pareto sorting.

ing MOP but different in treating the corresponding scalar (sigle-objective) optimization problems. NSGAII is without any doubt one of the most successful multi-objective genetic algorithm in the last decade. It introduced a nondominated sorting strategy, this strategy decreases the complexity of nondominated sorting from $O(MN^3)$ to $O(MN^2)$ and proposes a good approach to balance nonelitism and diversity of obtained solutions. In the last decade, NSGAII gains a large amount of citations and applications for its robustness and efficiency in solving MOPs.

Test instances used in this subsection are SCH, FON, KUR from [19] and ZDT1∼4, ZDT6, DTLZ1, DTLZ2 from [96]. Codes for MOEA/D and NSGAII are taken from Yarpiz (www.yarpiz.com). As well as MOEA/D [96] and NSGAII [19], the population size N is set to be 100 for 2-objective test instances and 150 for 3-objective test instances. The maximal number of generations is set to be 50 for 2-objective problems and 100



(a) NSM

(b) SNSM

Figure 5.10: Problems KUR solved by NSM and SNSM.

for 3-objective ones. The comparison is respect to two factors: numerical performance and complexity. Numerical performance is observed through depicting Pareto frontiers obtained by different methods in one figure, and complexity is measured by number of function value evaluation and time consumption.

Figure 5.11 depicts obtained approximate Pareto frontiers of 2-objectives Problems SCH, FON, KUR, ZDT1∼4 and ZDT6 solved by SNSM, MOEA/D and NSGAII, respectively. For Problem SCH (see Figure 5.11(a)), all the algorithms reach the real Pareto frontier, the diversity of solutions obtained by NSGAII and SNSM is better than that of MOEA/D. For Problem FON (see Figure 5.11(b)), Pareto frontier obtained by SNSM is much better than that obtained by MOGA/D and NSGAII not only in elitism but also in diversity. For Problem KUR (see Figure 5.11(c)), solutions obtained by SNSM perfectly simulated the disconnected Pareto frontier, while MOEA/D is not good at elitism and NSGAII only concentrate its solutions in the middle section. For Problem ZDT1 (see Figure 5.11(d)), all three algorithms performs evenly in elitism, but SNSM and NSGAII are better than MOEA/D in diversity. For Problem ZDT2 (see Figure 5.11(e)), MOEA/D and SNSM are neck and neck both in diversity and elitism, but both perform better than NSGAII. For Problem ZDT3 (see Figure 5.11(f)), although MOEA/D and SNSM perform better than NSGAII in elitism, solutions obtained by NSGAII simulate the Pareto frontier more comprehensively, and solutions obtained by SNSM are extremely dense in some area. For Problem ZDT4 (see Figure 5.11(g)), SNSM obtains solutions with better elitism, but MOEA/D and NSGAII obtain solutions with better diversity and distributed more comprehensively. For Problem ZDT6 (see Figure 5.11(h)), NSGAII performs better than both SNSM and MOEA/D not only in elitism but also in diversity, and again solutions obtained by SNSM concentrate in some points.

Figure 5.12 demonstrates numerical results of Problem DTLZ1 and DTLZ2 solved by

MOEA/D, NSGAII and SNSM, respectively. From Figure 5.12(a), 5.12(b) and 5.12(c), MOEA/D and SNSM obtain solutions which can normally simulate the real Pareto frontier, while NSGAII almost fail to solve this problem. Comparing with MOEA/D and SNSM, one can observe that solutions obtained by SNSM are better in both diversity and elitism than MOEA/D. From Figure 5.12(d), 5.12(e) and 5.12(f), it is obviously to observe that SNSM performs better than MOEA/D and NSGAII both in elitism and diversity.

Histogram 5.13 illustrates the complex comparison of SNSM, NSGAII and MOEAD for solving these problems. For 2-objective problems, the number of function evaluations of SNSM for most problems is larger than NSGAII and MOEAD, but it consumes less time than the other two for all problems. This reveals that SNSM is more efficient than the other two respect to number of function evaluations. For 3-objective problems, the number of function evaluations of SNSM is larger than the other two algorithm, its time consumption is only slightly larger. Actually, the reason for the number of function evaluations being large is that we used genetic algorithm to solve the scalar problems for Problem ZDT1$\sim$ZDT4, ZDT6, DTLZ1 and DTLZ2, if one can substitute genetic algorithm by other more efficient optimization solvers, the number of function evaluations could decrease dramatically.

To be summarized from Figure 5.11, 5.12 and Histogram 5.13, Algorithm SNSM (or EWSM for convex problems) could performs not worse and even better than MOEA/D and NSGAII. One reason for this advantage of SNSM is that its subproblem (Problem (5.3) or (5.6)) is solved using deterministic methods which is normally more accurate and faster than metaheuristic methods, and for regular problems, uniformly generated parameters (λ for EWSM or θ for SNSM) usually yield diversely distributed Pareto frontier. One may think that globally solving the subproblem is already a difficult task, let alone there are many subproblems need to be solved in SNSM. It is true that the problem is not suitable to be solved by SNSM if its subproblems are difficult to be globally solved. Another doubt about SNSM is that it should be inefficient because of these time consuming subproblems. This is not the case, because subproblem of SNSM are solve using deterministic methods which are quite efficient, so it will not take to much time for every subproblems. In fact, according to our numerical tests, the time consumed by solving a single subproblem of SNSM is generally less than that consumed by a generation of MOEA/D and NSGAII.

5.7.3 Numerical test and comparisons using CEC'09

In this subsection, we compare the numerical performance of SNSM with the methods proposed in the special session on performance assessment of unconstrained/bound constrained multi-objective optimization algorithms at CEC'09. There are 13 algorithms submitted to the special session:

(1) MOEAD [97];

(2) GDE3 [42];

(3) MOEADGM [9];

(4) MTS [82];

(5) LiuLiAlgorithm [50];

(6) DMOEADD [52];

(7) NSGAIILS [74];

(8) OWMOSaDE [35]

(9) ClusteringMOEA [83]

(10) AMGA [80]

(11) MOEP [66]

(12) DECMOSA-SQP [92];

(13) OMOEAII [27].

Figures 5.14 and 5.15 illustrate objective function value sets and real Pareto frontiers of these test problems (the figure of Problem 2 is skipped here since it is similar to Problem 1). Among these test problems, Problems 1-7 have two objective functions, whereas Problems 8-10 have three objective functions. The Pareto solutions of Problems 5, 6 and 9 are disconnected, while the others are connected.

In order to evaluate the numerical performance, we use the performance metric IGD. Suppose that P^* is a set of uniformly distributed points along the Pareto frontier. Let A be a set of solutions obtained by a certain solver. Then, the average distance from P^* to A is defined as

$$IGD(A, P^*) = \frac{\sum_{v \in P^*} d(v, A)}{|P^*|},$$

where $d(v, A)$ is the minimum Euclidean distance between v and the points in A, i.e.,

$$d(v, A) = \min_{y \in A} \parallel v - y \parallel .$$

In fact, P^* represents a sample set of the real Pareto frontier, if $|P^*|$ is large enough to approximate the Pareto frontier very well, $IGD(A, P^*)$ could measure both the diversity and convergence of A. A smaller $IGD(A, P^*)$ means the set A is closer to the real Pareto frontier and has better diversity.

In order to keep consistent with the final report of CEC'09 [95], in the implementation of SNSM, we compute 100 efficient solutions for problems with two objectives and 150 for problems with three objectives, the number of function evaluations is less than $300,000$. The numerical performance evaluated by IGD are illustrated in Table 5.2.

From Table 5.2, for Problems 1, 2 and 3, IGD evaluations rank at first, fifth and first, respectively. This means that SNSM performs better than other algorithms in solving Problems 1 and 3. When solving Problem 2, although the IGD evaluation of SNSM ranks at fifth, the difference with the first four algorithms are tiny, only in a precision of 10^{-3}. Additionally, the accuracy of IGD evaluations, which are 10^{-2}, reveals that Problems 1, 2 and 3 are prefect solved by SNSM. Figure 5.16 illustrates the obtained Pareto frontier of Problems 1 and 3, comparing with the real Pareto frontiers illustrated in Figure 5.15, we can conclude that these two problems are perfectly solved by SNSM.

For Problem 4, the IGD evaluation of SNSM ranks at the first, better than other referential algorithm. But its value is 0.01833 , only in a precision of 0.1, which is not perfect good. This point is also illustrated in Figure 5.17(a), which shows that the obtained efficient points are not extremely accurate and uniformly distributed.

SNSM is failed at solving Problem 5, one possible reason is that the Pareto frontier of Problem 5 consists of some isolated points, which is not suitable for SNSM.

The Pareto frontier of Problem 6 is disconnected, consist of two line segments and a point (See Figure 5.15). Figure 5.17(b) demonstrate the obtained Pareto frontier using SNSM, from the figure, we can observe that the real Pareto frontier is well simulated, points at the Pareto frontier have high accuracy and distribute evenly. In fact, from Table 5.2, IGD evaluation of SNSM for Problem 6 is 0.00976, accurate to 10^{-2}, ranks at the second.

For Problem 7 whose obtained Pareto frontier is illustrated in Figure 5.18(a), the IGD evaluation is 0.1063, ranks at the fifth. One interesting phenomenon showed in Figure 5.18(a) is that the lower part of the obtained Pareto frontier is very regular, but the upper part looks disorder. This may relate to the objective functions and option of θ.

For Problem 8, 9 and 10, IGD evaluation for SNSM are 0.1707, 0.03393 and 0.2382, rank at the seventh, first and second, respectively. The obtained Pareto frontier of Problem 9 is presented in Figure 5.18(b), Problem 8 and 10 are not presented since they are almost failed to simulate the real Pareto frontier. It is not uncommon that IGD evaluation for these three problems are not as small as others, because they all have three objective functions, which makes their Pareto frontiers surfaces. This not only increases the complexity of objective function of Problem (5.6), but also dramatically increases the amount of calculation since we have to work on much more points. In fact, even for the best MOEA, like MOEAD for Problem 8 (0.0584), DMOEADD for Problem 9 (0.04896) and MTS for Problem 10 (0.15306), the IGD evaluation is far from good enough.

Table 5.2: The numerical performance evaluated by IGD.

rank	UF1		UF2		UF3	
1	**SNSM**	**0.00381**	MTS	0.00615	**SNSM**	**0.00380**
2	MOEAD	0.00435	MOEADGM	0.0064	MOEAD	0.00742
3	GDE3	0.00534	DMOEADD	0.00679	LiuliAlgorithm	0.01497
4	MOEADGM	0.0062	MOEAD	0.00679	DMOEADD	0.03337
5	MTS	0.00646	**SNSM**	**0.0072**	MOEADGM	0.049
6	LiuLiAlgorithm	0.00785	OWMOSaDE	0.0081	MTS	0.0531
7	DMOEADD	0.01038	GDE3	0.01195	ClusteringMOEA	0.0549
8	NSGAIILS	0.01153	LiuLiAlgorithm	0.0123	AMGA	0.06998
9	OWMOSaDE	0.0122	NSGAIILS	0.01237	DECMOSA-SQP	0.0935
10	ClusteringMOEA	0.0299	AMGA	0.01623	MOEP	0.099
11	AMGA	0.03588	MOEP	0.0189	OWMOSaDE	0.103
12	MOEP	0.0596	ClusteringMOEA	0.0228	NSGAIILS	0.10603
13	DECMOSA-SQP	0.07702	DECMOSA-SQP	0.02834	GDE3	0.10639
14	OMOEAII	0.08564	OMOEAII	0.03057	OMOEAII	0.27141
rank	UF4		UF5		UF6	
1	**SNSM**	**0.01833**	MTS	0.01489	MOEAD	0.00587
2	MTS	0.02356	GDE3	0.03928	**SNSM**	**0.00976**
3	GDE3	0.0265	AMGA	0.09405	MTS	0.05917
4	DECMOSA-SQP	0.03392	LiuLiAlgorithm	0.16186	DMOEADD	0.06673
5	AMGA	0.04062	DECMOSA-SQP	0.16713	OMOEAII	0.07338
6	DMOEADD	0.04268	OMOEAII	0.1692	ClusteringMOEA	0.0871
7	MOEP	0.0427	MOEAD	0.18071	MOEP	0.1031
8	LiuLiAlgorithm	0.0435	MOEP	0.2245	DECMOSA-SQP	0.12604
9	OMOEAII	0.04624	ClusteringMOEA	0.2473	AMGA	0.12942
10	MOEADGM	0.0476	DMOEADD	0.31454	LiuLiAlgorithm	0.17555
11	OWMOSaDE	0.0513	OWMOSaDE	0.4303	OWMOSaDE	0.1918
12	NSGAIILS	0.0584	NSGAIILS	0.5657	GDE3	0.25091
13	ClusteringMOEA	0.0585	**SNSM**	**0.7032**	NSGAIILS	0.31032
14	MOEAD	0.06385	MOEADGM	1.7919	MOEADGM	0.5563
rank	UF7		UF8		UF9	
1	MOEAD	0.00444	MOEAD	0.0584	**SNSM**	**0.0339**
2	LiuLiAlgorithm	0.0073	DMOEADD	0.06841	DMOEADD	0.04896
3	MOEADGM	0.0076	LiuLiAlgorithm	0.08235	NSGAIILS	0.0719
4	DMOEADD	0.01032	NSGAIILS	0.0863	MOEAD	0.07896
5	**SNSM**	**0.01063**	OWMOSaDE	0.0945	GDE3	0.08248
6	MOEP	0.0197	MTS	0.11251	LiuliAlgorithm	0.09391
7	NSGAIILS	0.02132	**SNSM**	**0.1707**	OWMOSaDE	0.0983
8	ClusteringMOEA	0.0223	AMGA	0.17125	MTS	0.11442
9	DECMOSA-SQP	0.02416	OMOEAII	0.192	DECMOSA-SQP	0.14111
10	GDE3	0.02522	DECMOSA-SQP	0.21583	MOEADGM	0.1878
11	OMOEAII	0.03354	ClusteringMOEA	0.2383	AMGA	0.18861
12	MTS	0.04079	MOEADGM	0.2446	OMOEAII	0.23179
13	AMGA	0.05707	GDE3	0.24855	ClusteringMOEA	0.2934
14	OWMOSaDE	0.0585	MOEP	0.423	MOEP	0.342
rank	UF10					
1	MTS	0.15306				
2	**SNSM**	**0.2382**				
3	DMOEADD	0.32211				
4	AMGA	0.32418				
5	MOEP	0.3621				
6	DECMOSA-SQP	0.36985				
7	ClusteringMOEA	0.4111				
8	GDE3	0.43326				
9	LiuLiAlgorithm	0.44691				
10	MOEAD	0.47415				
11	MOEADGM	0.5646				
12	OMOEAII	0.62754				
13	OWMOSaDE	0.743				
14	NSGAIILS	0.84468				

5.7.4 Solving the mean-variance-skewness model

In the above section, we have tested our algorithm through benchmarks. In this part, we will apply the algorithm to solve multi-objective portfolio selection problem. 10 portfolios in steel industry from Shanghai stock exchange are selected. Since the skewness is non-concave, the Pareto-front of this tri-objective optimisation problem is hard to be approximated through linear weighting methods. Now SNSM is applied to solve return-risk-skewness portfolio selection problem. The distributions of Pareton-optimal set over return-risk and return-risk-skewness are shown in Fig. 5.19 and Fig. 5.20.

To verify the effectiveness of our method, the results are compared with MOEAD and NSGAIILS. However, the metric $IGD(A, P^*)$ cannot be used to assess the performance of different algorithms as P^* is unknown. In MOEAD and NSGAIILS, the constraints $\sum_{=1}^{n} x_i = 1$ is transformed through the following reweighting method in each iteration:

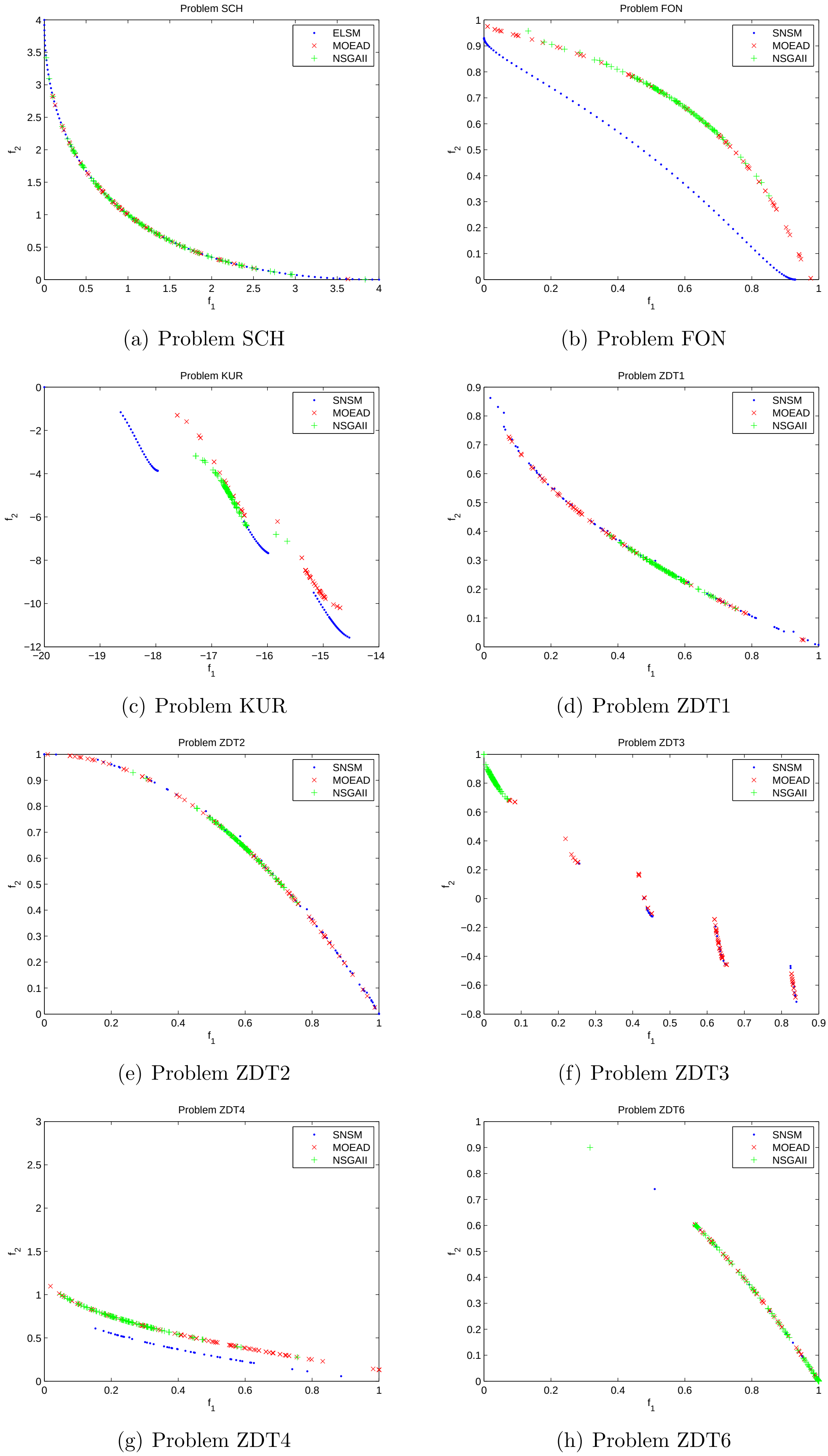
$$x_i = \frac{x_i}{\sum_i x_i}, \forall i \in D,$$

where D is the set of all genetics or particles used in each iteration. In this comparison, we select 150 points in SNSM and 100 genetics in MOEAD and NSGAIILS. The solutions obtained by the three methods are depicted in Fig. 5.21. From Fig. 5.21, we can see that all three methods can approximate Pareto-front and no significant gaps are observed for the three method. The reason behind might be that the Pareto-front for the problem (5.10) looks like convex and thus all of the three methods can approximate Pareto-front perfectly.

5.8 Conclusion

This chapter proposed population-based linear and nonlinear scalarization methods for MOPs which is applied to solve mean-variance-skewness portfolio selection. Scalarization is an important type of strategy to handle MOPs. The previous research mainly focuses on scalar techniques, while this paper contributes to generalizing the scalarization methods to population-based ones. We first extended the weighted sum method to a population-based case which has good theoretical properties and numerical performances for convex MOPs, but fails to solve nonconvex MOPs. In order to handle nonconvex MOPs, we designed a nonlinear scalar technique which transforms an MOP to a nonlinear scalar optimisation problem. It can be proved that, in some conditions, the global optimal solution of the nonlinear optimisation problem must be an efficient solution of the original multi-objective problem. Based on this property, a nonlinear scalarization method and a slack variation of it were proposed. A wide range of numerical tests were presented. First of all, numerical performance of the proposed methods were illustrated by some academic multi-objective optimisation benchmarks; then numerical comparisons among the proposed methods and two typical multi-objective optimisation methods MOEA/D and NSGAII were made; finally the proposed method were applied to solve CEC'09 test instances and the results were compared with 13 referential algorithms proposed in CEC'09. Numerical tests show that methods proposed in this paper are able to solve MOPs with promising elitism and diversity. The mean-varaince-skewness portfolio selection problem is solved through our proposed nonlinear scalarisation method since the objective skewness is non-convex.

There are two critical points that need to be tackled in our future work on this subject. First, a global optimisation method plays a very important role in the proposed methods with a fast deterministic global optimisation method dramatically increasing efficiency. In this paper, subproblems are solved directly using a global optimisation method, but for MOPs whose objective functions are extremely complicated, this strategy may not work. Therefore, some metaheuristic strategies should be introduced to handle subproblems. Second, the distribution of predetermined scalarization parameters corresponds to the distribution of obtained solutions. For regular problems, uniformly generated scalarization parameters usually yields a diverse distribution of solutions, but for irregular problems, many scalarization parameters may correspond with one solution which destroys the diversity of the obtained solutions. In this case, a self-adaptive strategy for generating scalarization parameters should be developed.



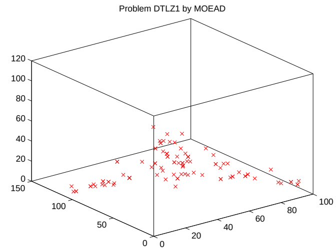
(a) Problem SCH (b) Problem FON

(c) Problem KUR (d) Problem ZDT1

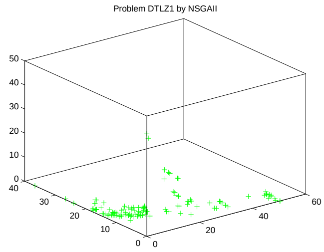
(e) Problem ZDT2 (f) Problem ZDT3

(g) Problem ZDT4 (h) Problem ZDT6

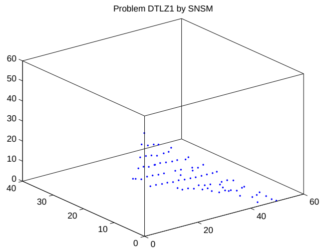
Figure 5.11: Numerical performance for 2-objective problems.



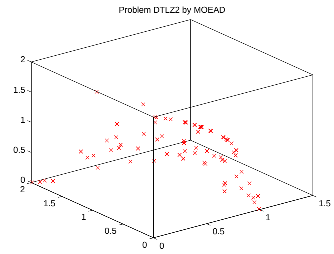
(a) Problem DTLZ1 by MOEAD



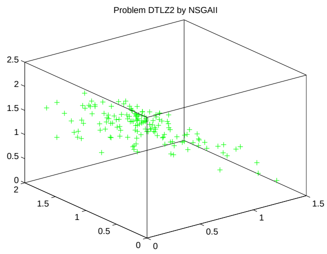
(b) Problem DTLZ1 by NSGAII



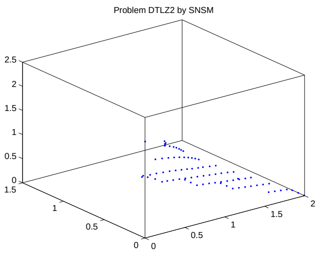
(c) Problem DTLZ1 by SNSM



(d) Problem DTLZ2 by MOEAD

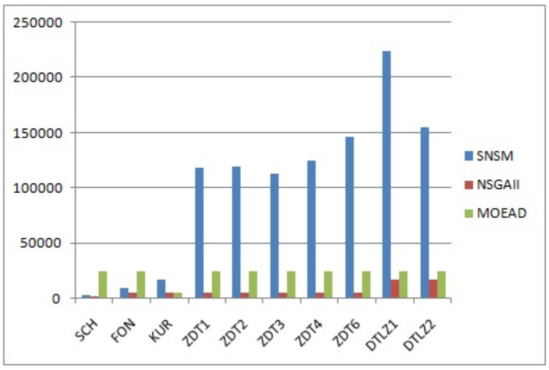


(e) Problem DTLZ2 by NSGAII

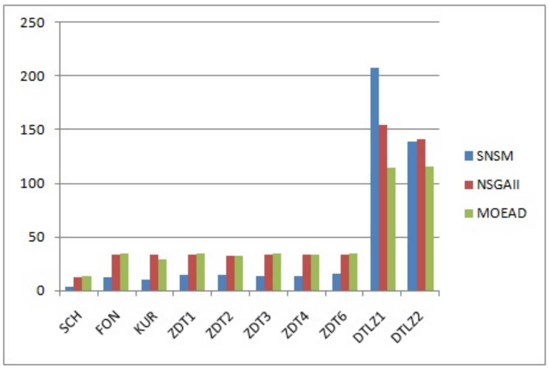


(f) Problem DTLZ2 by SNSM

Figure 5.12: Numerical performance for 3-objective problems.



(a) Number of function evaluation



(b) Time consumption

Figure 5.13: Comparison respect to time consumption and number of function value evaluation.

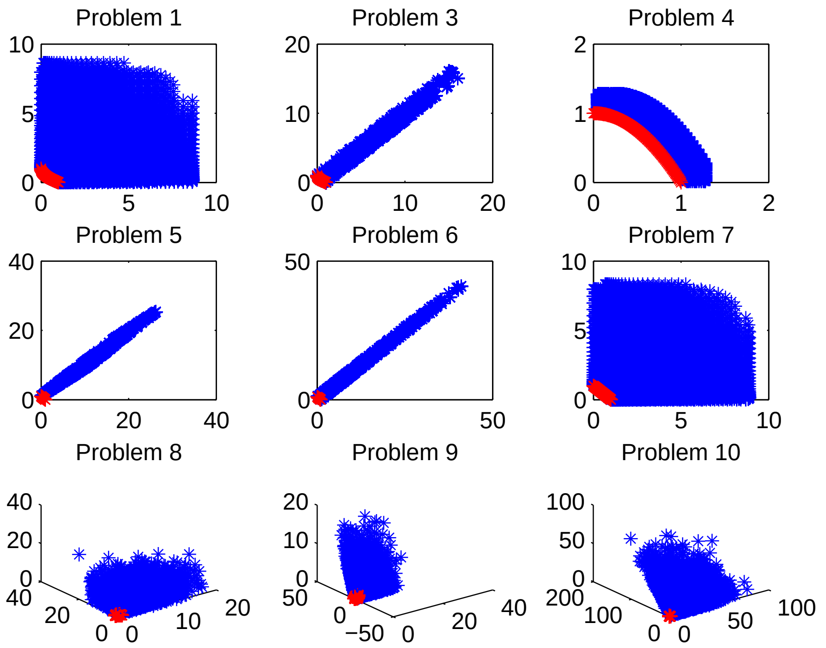


Figure 5.14: Objective function value space for test problems.

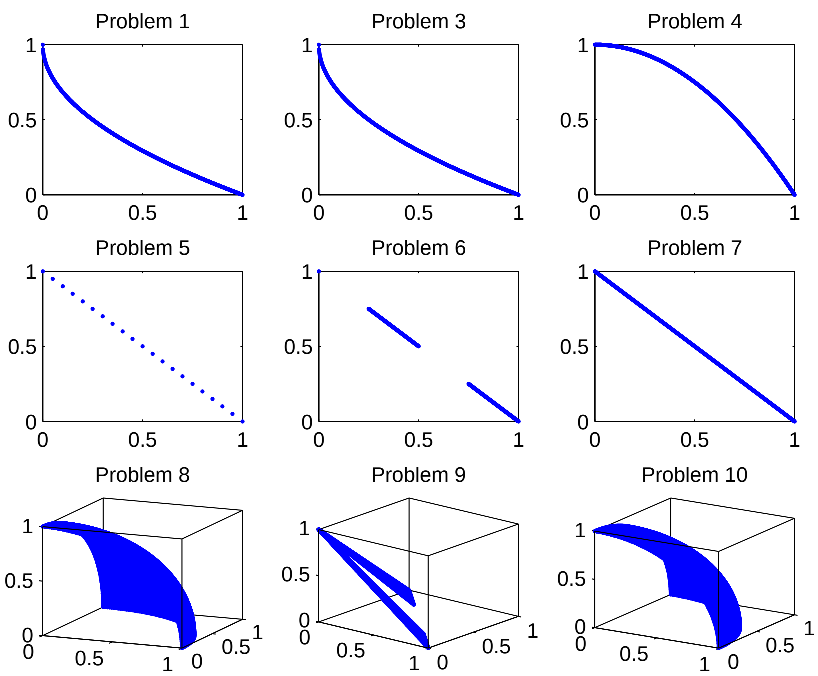
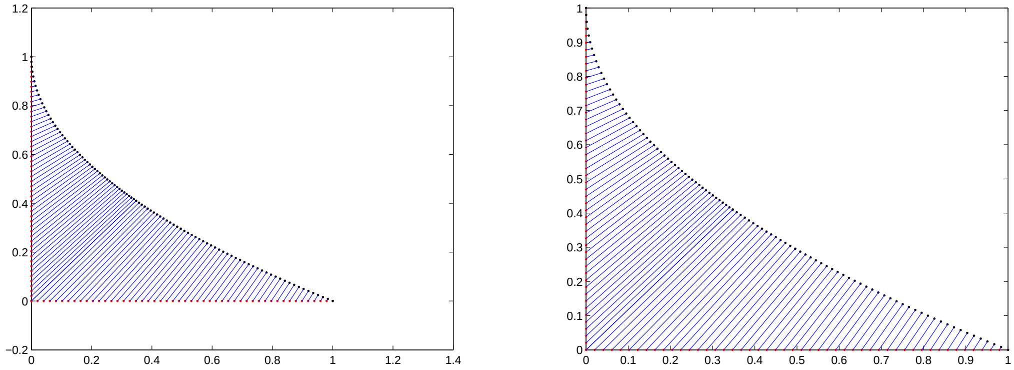
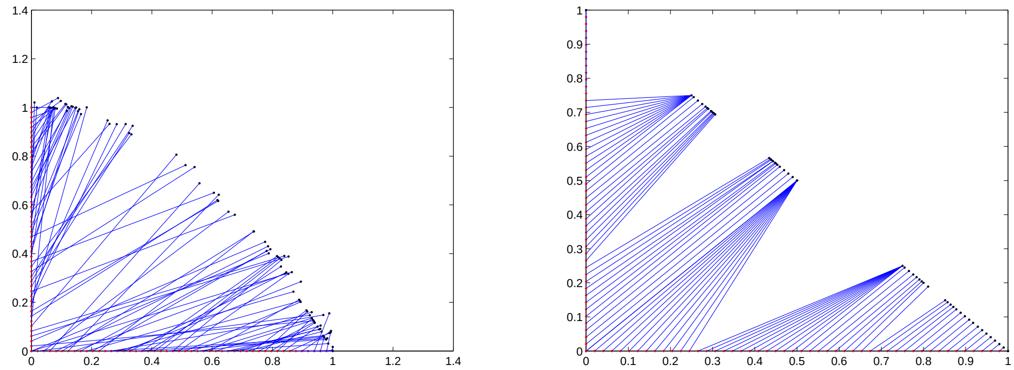


Figure 5.15: Real Pareto frontiers for test problems.



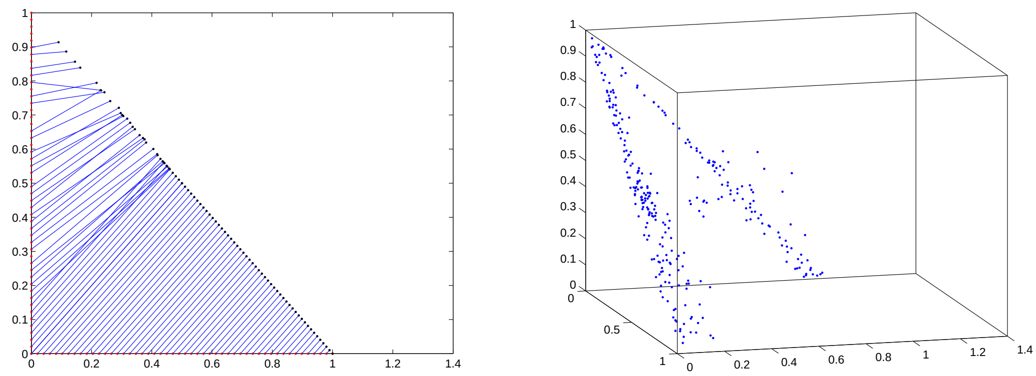
(a) Problem 1 (b) Problem 3

Figure 5.16: Solving Problem 1 and 3 using SNSM.



(a) Problem 4 (b) Problem 6

Figure 5.17: Solving Problem 4 and 6 using SNSM.



(a) Problem 7 (b) Problem 9

Figure 5.18: Solving Problem 7 and 9 using SNSM.

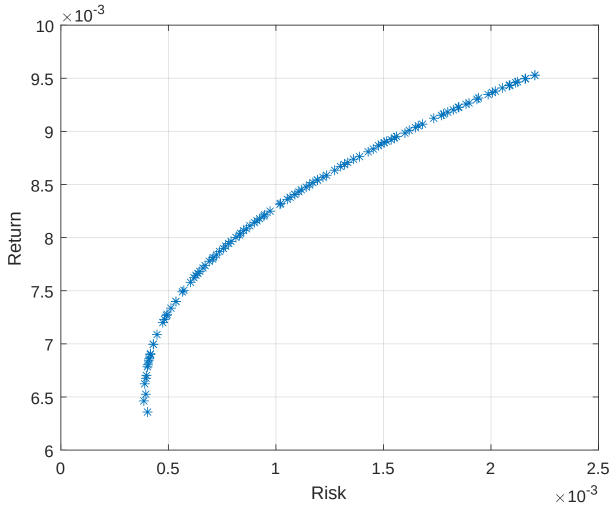


Figure 5.19: Return-Risk Pareto-front.

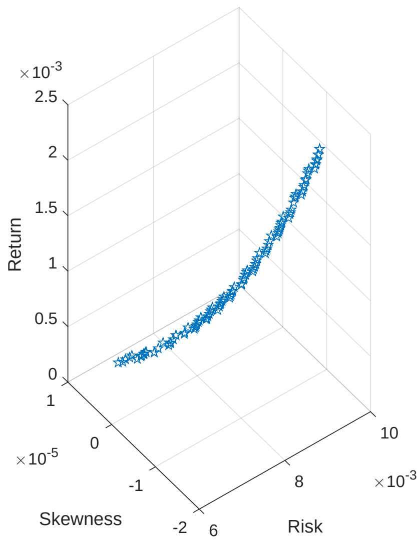


Figure 5.20: Return-Risk Pareto front.

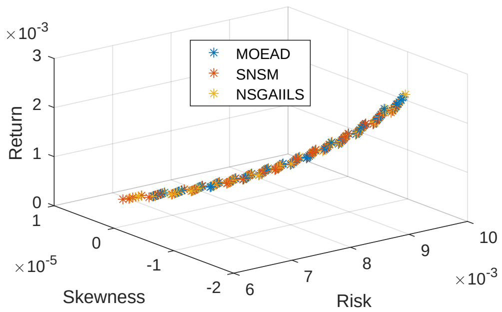


Figure 5.21: The Pareto-frontiers obtained by SNSM, MOEAD and NSGAIILS.

CHAPTER 6

Conclusions and Future Research

6.1 Main Contributions of this Thesis

In this thesis, we consider portfolio selection problems under three different scenarios. For each of problem, we formulate it as either robust optimisation problem or multi-objective optimisation problem. Then, the corresponding numerical algorithms are developed to solve them. The main contributions are summarised as follows:

In Chapter 2, the existing results on portfolio selection is reviewed. Chapter 3 s-tudied portfolio selection problem with distribution and moment uncertainty subject to bankruptcy constraints. If the distribution is uncertain but with exact moments, we prove that it can be equivalently transformed into a second-order-cone programming problem. If the moments are also uncertain, we further improve that it can be transformed into an equivalent second-order-cone programming problem. Numerical examples are used to illustrate our proposed method.

Chapter 4 studied a multi-period multi-objective portfolio selection problem and the moments are also uncertain, but within an elliptical uncertainty set. This problem has been formulated as a min-max optimisation problem. Through analysing the model theoretically, we showed that it can be transformed a simple min-max optimisation problem, where the inner maximisation problem is 1-dimensional and concave. Thus, the inner maximisation problem was solved semi-analytically as the optimal solution can only be achieved at the boundary or the inner equilibrium point. Thus, the original problem was easily solved through the existing convex software.

Chapter 5 proposed a nonlinear scarlarisation method for multi-objective optimisation problem. Most of exact-based multi-objective methods transform a multi-objective optimisation problem into a single-objective optimisation problem through linear weighting method. If the original single-objective optimisation problem is convex, this transformation can explore all the solutions in the Pareto front. However, if the original single-objective optimisation problem is not convex, then the Pareto front might not be approximated efficiently through linear scarlarisation and nonlinear scarlarisation might provide better approximation. Our method seems be the first one to transform a multi-objective optimsiation problem into a single-objective optimisation problem. Numerical performances on test problems are promising. We applied this method to solve mean-variance-skewness portfolio selection problem and the Pareto-front was efficiently approximated.

6.2 Future Research Directions

Portfolio selection is a very complicated problem due to the complexness of the financial markets and investor unpredictability. Although it has been over 70 years since the first seminal work on portfolio selection in 1950s, this problem is far from being solved. How to address the uncertainty of the financial markets is still challenging. In Chapter 3, we studied multi-period portfolio selection with distributional uncertainty. The problem that we studied is still simple where only chance constraints are involved. If the risk is measured by VaR or CVaR and more constraints from financial markets and investment requirements, how to solve this problem is still waiting for investigation. In Chapter 4, the uncertainty set is modelled as an elliptical set. How to determine this set efficiently and if the risk is measured by other risk measurements, how to solve the problem is still not resolved. In Chapter 5, our tri-objective portfolio selection problem has included only simple constraints. If more complicated constraints are involved, how to solve this multi-objective optimisation problem is also worth for further investigation.

Bibliography

[1] F. B. Abdelaziz, B. Aouni, and R. El Fayedh, "Multi-objective stochastic programming for portfolio selection," *European Journal of Operational Research*, vol. 177, no. 3, pp. 1811–1823, 2007.

[2] G. J. Alexander and A. M. Baptista, "Economic implications of using a mean-var model for portfolio selection: A comparison with mean-variance analysis," *Journal of Economic Dynamics and Control*, vol. 26, no. 7-8, pp. 1159–1193, 2002.

[3] P. Artzner, F. Delbaen, J.-M. Eber, and D. Heath, "Coherent measures of risk," *Mathematical finance*, vol. 9, no. 3, pp. 203–228, 1999.

[4] P. Artzner, F. Delbaen, J.-M. Eber, D. Heath, and H. Ku, "Coherent multiperiod risk adjusted values and bellmans principle," *Annals of Operations Research*, vol. 152, no. 1, pp. 5–22, 2007.

[5] G. Beliakov and A. Bagirov, "Non-smooth optimization methods for computation of the conditional value-at-risk and portfolio optimization," *Optimization*, vol. 55, no. 5-6, pp. 459–479, 2006.

[6] A. Ben-Tal and A. Nemirovski, "Robust optimization–methodology and applications," *Mathematical Programming*, vol. 92, no. 3, pp. 453–480, 2002.

[7] L. Bin, "A new risk measure and its application in portfolio optimization: The spp–cvar approach," *Economic Modelling*, vol. 51, pp. 383–390, 2015.

[8] R. P. Brito and L. N. Vicente, "Efficient cardinality/mean-variance portfolios," in *IFIP Conference on System Modeling and Optimization*. Springer, 2013, pp. 52–73.

[9] C.-M. Chen, Y.-p. Chen, and Q. Zhang, "Enhancing moea/d with guided mutation and priority update for multi-objective optimization," in *2009 IEEE Congress on Evolutionary Computation*. IEEE, 2009, pp. 209–216.

[10] L. Chen, S. He, and S. Zhang, "Tight bounds for some risk measures, with applications to robust portfolio selection," *Operations Research*, vol. 59, no. 4, pp. 847–865, 2011.

[11] X. Cui, J. Gao, X. Li, and D. Li, “Optimal multi-period mean–variance policy under no-shorting constraint,” *European Journal of Operational Research*, vol. 234, no. 2, pp. 459–468, 2014.

[12] X. Cui, J. Gao, and Y. Shi, “Multi-period mean–variance portfolio optimization with management fees,” *Operational Research*, pp. 1–22, 2019.

[13] X. Cui, J. Gao, Y. Shi, and S. Zhu, “Time-consistent and self-coordination strategies for multi-period mean-conditional value-at-risk portfolio selection,” *European Journal of Operational Research*, pp. 781–789, 2019.

[14] X. Cui, X. Li, and D. Li, “Unified framework of mean-field formulations for optimal multi-period mean-variance portfolio selection,” *IEEE transactions on automatic control*, vol. 59, no. 7, pp. 1833–1844, 2014.

[15] X. Cui and Y. Shi, “Multiperiod mean-cvar portfolio selection,” in *Modelling, Computation and Optimization in Information Systems and Management Sciences.* Springer, 2015, pp. 293–304.

[16] X. Cui, X. Sun, and D. Sha, “An empirical study on discrete optimization models for portfolio selection,” *Journal of Industrial and Management Optimization*, vol. 5, pp. 33–46, 2009.

[17] G. B. Dantzig and G. Infanger, “Multi-stage stochastic linear programs for portfolio optimization,” *Annals of Operations Research*, vol. 45, no. 1, pp. 59–76, 1993.

[18] I. Das and J. E. Dennis, “Normal-boundary intersection: A new method for generating the pareto surface in nonlinear multicriteria optimization problems,” *SIAM journal on optimization*, vol. 8, no. 3, pp. 631–657, 1998.

[19] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan, “A fast and elitist multiobjective genetic algorithm: Nsga-ii,” *IEEE transactions on evolutionary computation*, vol. 6, no. 2, pp. 182–197, 2002.

[20] E. Delage and Y. Ye, “Distributionally robust optimization under moment uncertainty with application to data-driven problems,” *Operations research*, vol. 58, no. 3, pp. 595–612, 2010.

[21] M. Ehrgott, *Multicriteria optimization.* Springer Science & Business Media, 2005, vol. 491.

[22] G. Eichfelder, “An adaptive scalarization method in multiobjective optimization,” *SIAM Journal on Optimization*, vol. 19, no. 4, pp. 1694–1718, 2009.

[23] J. Estrada, "Mean-semivariance optimization: A heuristic approach," *Journal of Applied Finance (Formerly Financial Practice and Education)*, vol. 18, no. 1, 2008.

[24] H. Fazlollahtabar and M. Saidi-Mehrabad, "Optimizing multi-objective decision making having qualitative evaluation," *Journal of Industrial and Management Optimization*, vol. 11, no. 3, pp. 747–762, 2015.

[25] J. Gao and D. Li, "Optimal cardinality constrained portfolio selection," *Operations research*, vol. 61, no. 3, pp. 745–761, 2013.

[26] J. Gao, D. Li, X. Cui, and S. Wang, "Time cardinality constrained mean–variance dynamic portfolio selection and market timing: A stochastic control approach," *Automatica*, vol. 54, pp. 91–99, 2015.

[27] S. Gao, B. Xiao, L. Zhang, Y. Shi, X. Tian, Y. Yang, H. Long, X. Yang, D. Yu, and Z. Yan, "Differential evolution with self-adaptation and local search for constrained multiobjective optimization," in *Proceedings of the IEEE Congress on Evolutionary Computation, CEC 2009, Trondheim, Norway, 1959-1964 May, 2009*, 2009.

[28] A. M. Geoffrion, "Proper efficiency and the theory of vector maximization," *Journal of mathematical analysis and applications*, vol. 22, no. 3, pp. 618–630, 1968.

[29] A. Ghahtarani and A. A. Najafi, "Robust goal programming for multi-objective portfolio selection problem," *Economic Modelling*, vol. 33, pp. 588–592, 2013.

[30] L. E. Ghaoui, M. Oks, and F. Oustry, "Worst-case value-at-risk and robust portfolio optimization: A conic programming approach," *Operations research*, vol. 51, no. 4, pp. 543–556, 2003.

[31] D. Goldfarb and G. Iyengar, "Robust portfolio selection problems," *Mathematics of operations research*, vol. 28, no. 1, pp. 1–38, 2003.

[32] A. Goli, H. K. Zare, R. Tavakkoli-Moghaddam, and A. Sadeghieh, "Hybrid artificial intelligence and robust optimization for a multi-objective product portfolio problem case study: The dairy products industry," *Computers & Industrial Engineering*, vol. 137, p. 106090, 2019.

[33] R. Hartley, "On cone-efficiency, cone-convexity and cone-compactness," *SIAM Journal on Applied Mathematics*, vol. 34, no. 2, pp. 211–222, 1978.

[34] D. Huang, S. Zhu, F. J. Fabozzi, and M. Fukushima, "Portfolio selection under distributional uncertainty: A relative robust cvar approach," *European Journal of Operational Research*, vol. 203, no. 1, pp. 185–194, 2010.

[35] V. L. Huang, S. Z. Zhao, R. Mallipeddi, and P. N. Suganthan, "Multi-objective optimization using self-adaptive differential evolution algorithm," in *2009 IEEE Congress on Evolutionary Computation*. IEEE, 2009, pp. 190–194.

[36] P. Jana, T. Roy, and S. Mazumder, "Multi-objective possibilistic model for portfolio selection with transaction cost," *Journal of computational and applied mathematics*, vol. 228, no. 1, pp. 188–196, 2009.

[37] X. Jin, N. Chen, and Y. Yuan, "Multi-period and tri-objective uncertain portfolio selection model: A behavioral approach," *The North American Journal of Economics and Finance*, vol. 47, pp. 492–504, 2019.

[38] C. B. Kalayci, O. Ertenlice, H. Akyer, and H. Aygoren, "An artificial bee colony algorithm with feasibility enforcement and infeasibility toleration procedures for cardinality constrained portfolio optimization," *Expert Systems with Applications*, vol. 85, pp. 61–75, 2017.

[39] C. B. Kalayci, O. Polat, and M. A. Akbay, "An efficient hybrid metaheuristic algorithm for cardinality constrained portfolio optimization," *Swarm and Evolutionary Computation*, vol. 54, p. 100662, 2020.

[40] R. Kamali, S. Mahmoodi, and M.-T. Jahandideh, "Optimization of multi-period portfolio model after fitting best distribution," *Finance Research Letters*, vol. 30, pp. 44–50, 2019.

[41] K. Khalili-Damghani and M. Amiri, "Solving binary-state multi-objective reliability redundancy allocation series-parallel problem using efficient epsilon-constraint, multi-start partial bound enumeration algorithm, and dea," *Reliability Engineering & System Safety*, vol. 103, pp. 35–44, 2012.

[42] S. Kukkonen and J. Lampinen, "Performance assessment of generalized differential evolution 3 with a given set of constrained multi-objective test problems," in *2009 IEEE Congress on Evolutionary Computation*. IEEE, 2009, pp. 1943–1950.

[43] M. Laumanns, L. Thiele, and E. Zitzler, "An efficient, adaptive parameter variation scheme for metaheuristics based on the epsilon-constraint method," *European Journal of Operational Research*, vol. 169, no. 3, pp. 932–942, 2006.

[44] M. A. Lejeune and S. Shen, "Multi-objective probabilistically constrained programs with variable risk: Models for multi-portfolio financial optimization," *European Journal of Operational Research*, vol. 252, no. 2, pp. 522–539, 2016.

[45] B. Li, Y. Zhu, Y. Sun, G. Aw, and K. L. Teo, "Multi-period portfolio selection problem under uncertain environment with bankruptcy constraint," *Applied Mathematical Modelling*, vol. 56, pp. 539–550, 2018.

[46] D. Li and W.-L. Ng, "Optimal dynamic portfolio selection: Multiperiod mean-variance formulation," *Mathematical finance*, vol. 10, no. 3, pp. 387–406, 2000.

[47] K. Liagkouras and K. Metaxiotis, "A new efficiently encoded multiobjective algorithm for the solution of the cardinality constrained portfolio optimization problem," *Annals of Operations Research*, vol. 267, no. 1-2, pp. 281–319, 2018.

[48] C.-C. Lin and Y.-T. Liu, "Genetic algorithms for portfolio selection problems with minimum transaction lots," *European Journal of Operational Research*, vol. 185, no. 1, pp. 393–404, 2008.

[49] A. Ling, J. Sun, and M. Wang, "Robust multi-period portfolio selection based on downside risk with asymmetrically distributed uncertainty set," *European Journal of Operational Research*, 2019.

[50] H.-L. Liu and X. Li, "The multiobjective evolutionary algorithm based on determined weight and sub-regional search," in *2009 IEEE Congress on Evolutionary Computation*. IEEE, 2009, pp. 1928–1934.

[51] J. Liu, Z. Chen, A. Lisser, and Z. Xu, "Closed-form optimal portfolios of distributionally robust mean-cvar problems with unknown mean and variance," *Applied Mathematics & Optimization*, vol. 79, no. 3, pp. 671–693, 2019.

[52] M. Liu, X. Zou, Y. Chen, and Z. Wu, "Performance assessment of dmoea-dd with cec 2009 moea competition test instances," in *2009 IEEE Congress on Evolutionary Computation*. IEEE, 2009, pp. 2913–2918.

[53] Q. Long, "A constraint handling technique for constrained multi-objective genetic algorithm," *Swarm and Evolutionary Computation*, vol. 15, pp. 66–79, 2014.

[54] K. T. Lwin, R. Qu, and B. L. MacCarthy, "Mean-var portfolio optimization: A nonparametric approach," *European Journal of Operational Research*, vol. 260, no. 2, pp. 751–766, 2017.

[55] H. Markowitz, "Portfolio selection," *The journal of finance*, vol. 7, no. 1, pp. 77–91, 1952.

[56] R. T. Marler and J. S. Arora, "The weighted sum method for multi-objective optimization: new insights," *Structural and multidisciplinary optimization*, vol. 41, no. 6, pp. 853–862, 2010.

[57] S. S. Meghwani and M. Thakur, "Multi-criteria algorithms for portfolio optimization under practical constraints," *Swarm and evolutionary computation*, vol. 37, pp. 104–125, 2017.

[58] ——, "Multi-objective heuristic algorithms for practical portfolio optimization and rebalancing with transaction cost," *Applied Soft Computing*, vol. 67, pp. 865–894, 2018.

[59] R. C. Merton, "Optimum consumption and portfolio rules in a continuous-time model," in *Stochastic Optimization Models in Finance.* Elsevier, 1975, pp. 621–661.

[60] R. O. Michaud and R. O. Michaud, *Efficient asset management: a practical guide to stock portfolio optimization and asset allocation.* Oxford University Press, 2008.

[61] S. Mousavi, A. Esfahanipour, and M. H. Fazel Zarandi, "A modular takagi-sugeno-kang (tsk) system based on a modified hybrid soft clustering for stock selection," *Scientia Iranica*, 2019.

[62] M. Ç. Pınar, "Robust scenario optimization based on downside-risk measure for multi-period portfolio selection," *OR Spectrum*, vol. 29, no. 2, pp. 295–309, 2007.

[63] N. Pindoriya, S. Singh, and S. Singh, "Multi-objective mean–variance–skewness model for generation portfolio allocation in electricity markets," *Electric Power Systems Research*, vol. 80, no. 10, pp. 1314–1321, 2010.

[64] A. R. Pouya, M. Solimanpur, and M. J. Rezaee, "Solving multi-objective portfolio optimization problem using invasive weed optimization," *Swarm and Evolutionary Computation*, vol. 28, pp. 42–57, 2016.

[65] R. Qi and G. G. Yen, "Hybrid bi-objective portfolio optimization with pre-selection strategy," *Information Sciences*, vol. 417, pp. 401–419, 2017.

[66] B. Y. Qu, P. N. Suganthan, S. Member, and IEEE, "Multi-objective evolutionary programming without non-domination sorting is up to twenty times faster," in *Proceedings of the IEEE Congress on Evolutionary Computation, CEC 2009, Trondheim, Norway, 18-21 May, 2009*, 2009.

[67] A. Rigamonti, "Downside risk optimization vs mean-variance optimization," *Available at SSRN 3465717*, 2019.

[68] R. T. Rockafellar, S. Uryasev *et al.*, "Optimization of conditional value-at-risk," *Journal of risk*, vol. 2, pp. 21–42, 2000.

[69] K. Ruan and M. Fukushima, "Robust portfolio selection with a combined wcvar and factor model," *Journal of Industrial and Management Optimization*, vol. 8, no. 2, p. 343, 2012.

[70] R. Ruiz-Torrubiano and A. Suárez, "A memetic algorithm for cardinality-constrained portfolio optimization with transaction costs," *Applied Soft Computing*, vol. 36, pp. 125–142, 2015.

[71] K. Schöttle and R. Werner, "Robustness properties of mean-variance portfolios," *Optimization*, vol. 58, no. 6, pp. 641–663, 2009.

[72] M. Shaverdi, S. Yaghoubi, and H. Ensafian, "A multi-objective robust possibilistic model for technology portfolio optimization considering social impact and different types of financing," *Applied Soft Computing*, vol. 86, p. 105892, 2020.

[73] Y. L. T. Silva, A. B. Herthel, and A. Subramanian, "A multi-objective evolutionary algorithm for a class of mean-variance portfolio selection problems," *Expert Systems with Applications*, vol. 133, pp. 225–241, 2019.

[74] K. Sindhya, A. Sinha, K. Deb, and K. Miettinen, "Local search based evolutionary multi-objective optimization algorithm for constrained and unconstrained problems," in *2009 IEEE Congress on Evolutionary Computation*. IEEE, 2009, pp. 2919–2926.

[75] H. Soleimani, H. R. Golmakani, and M. H. Salimi, "Markowitz-based portfolio selection with minimum transaction lots, cardinality constraints and regarding sector capitalization using genetic algorithm," *Expert Systems with Applications*, vol. 36, no. 3, pp. 5058–5063, 2009.

[76] M. C. Steinbach, "Markowitz revisited: Mean-variance models in financial portfolio analysis," *SIAM review*, vol. 43, no. 1, pp. 31–85, 2001.

[77] M. S. Strub, D. Li, X. Cui, and J. Gao, "Discrete-time mean-cvar portfolio selection and time-consistency induced term structure of the cvar," *Journal of Economic Dynamics and Control*, vol. 108, p. 103751, 2019.

[78] Y. Sun, G. Aw, R. Loxton, and K. L. Teo, "Chance-constrained optimization for pension fund portfolios in the presence of default risk," *European Journal of Operational Research*, vol. 256, no. 1, pp. 205–214, 2017.

[79] Y. Tian, S. Fang, Z. Deng, and Q. Jin, "Cardinality constrained portfolio selection problem: A completely positive programming approach," *Journal of Industrial & Management Optimization*, vol. 12, no. 3, p. 1041, 2016.

[80] S. Tiwari, G. Fadel, P. Koch, and K. Deb, "Performance assessment of the hybrid archive-based micro genetic algorithm (amga) on the cec09 test problems," in *2009 IEEE Congress on Evolutionary Computation.* IEEE, 2009, pp. 1935–1942.

[81] X. Tong, L. Qi, F. Wu, and H. Zhou, "A smoothing method for solving portfolio optimization with cvar and applications in allocation of generation asset," *Applied Mathematics and Computation*, vol. 216, no. 6, pp. 1723–1740, 2010.

[82] L.-Y. Tseng and C. Chen, "Multiple trajectory search for unconstrained/constrained multi-objective optimization," in *2009 IEEE Congress on Evolutionary Computation.* IEEE, 2009, pp. 1951–1958.

[83] Y. Wang, C. Dang, H. Li, L. Han, and J. Wei, "A clustering multi-objective evolutionary algorithm based on orthogonal and uniform design," in *2009 IEEE congress on evolutionary computation.* IEEE, 2009, pp. 2927–2933.

[84] Z. Wang and S. Liu, "Multi-period mean-variance portfolio selection with fixed and proportional transaction costs," *Journal of industrial and management optimization*, vol. 9, no. 3, pp. 643–657, 2013.

[85] C. Wu, K. L. Teo, and S. Wu, "Min–max optimal control of linear systems with uncertainty and terminal state constraints," *Automatica*, vol. 49, no. 6, pp. 1809–1815, 2013.

[86] P. Xidonas, G. Mavrotas, C. Hassapis, and C. Zopounidis, "Robust multiobjective portfolio optimization: A minimax regret approach," *European Journal of Operational Research*, vol. 262, no. 1, pp. 299–305, 2017.

[87] W. Yan, R. Miao, and S. Li, "Multi-period semi-variance portfolio selection: Model and numerical solution," *Applied Mathematics and Computation*, vol. 194, no. 1, pp. 128–134, 2007.

[88] H. Yao, Z. Li, and Y. Lai, "Mean–cvar portfolio selection: A nonparametric estimation framework," *Computers & Operations Research*, vol. 40, no. 4, pp. 1014–1022, 2013.

[89] H. Yao, Z. Li, and D. Li, "Multi-period mean-variance portfolio selection with stochastic interest rate and uncontrollable liability," *European Journal of Operational Research*, vol. 252, no. 3, pp. 837–851, 2016.

[90] L. Yi, Z. Li, and D. Li, "Multi-period portfolio selection for asset-liability management with uncertain investment horizon," *Journal of industrial and management optimization*, vol. 4, no. 3, pp. 535–552, 2008.

[91] L. Yu, S. Wang, and K. K. Lai, "Neural network-based mean–variance–skewness model for portfolio selection," *Computers & Operations Research*, vol. 35, no. 1, pp. 34–46, 2008.

[92] A. Zamuda, J. Brest, B. Boskovic, and V. Zumer, "Differential evolution with self-adaptation and local search for constrained multiobjective optimization," in *Proceedings of the IEEE Congress on Evolutionary Computation, CEC 2009, Trondheim, Norway, 195-202 May, 2009*, 2009.

[93] N. Zhang, "A symmetric gauss-seidel based method for a class of multi-period mean-variance portfolio selection problems," *Journal of Industrial & Management Optimization*, vol. 13, no. 5, p. 1, 2018.

[94] P. Zhang, "Chance-constrained multiperiod mean absolute deviation uncertain portfolio selection," *Journal of Industrial & Management Optimization*, vol. 15, no. 2, pp. 537–564, 2019.

[95] Q. Zhang and P. Suganthan, "Final report on cec09 moea competition, technical report," 2009.

[96] Q. Zhang and H. Li, "Moea/d: A multiobjective evolutionary algorithm based on decomposition," *IEEE Transactions on evolutionary computation*, vol. 11, no. 6, pp. 712–731, 2007.

[97] Q. Zhang, W. Liu, and H. Li, "The performance of a new version of moea/d on cec09 unconstrained mop test instances," in *2009 IEEE congress on evolutionary computation*. IEEE, 2009, pp. 203–208.

[98] Y. Zhang, R. Jiang, and S. Shen, "Ambiguous chance-constrained binary programs under mean-covariance information," *SIAM Journal on Optimization*, vol. 28, no. 4, pp. 2922–2944, 2018.

[99] C. Zhao, C. Wu, J. Chai, X. Wang, X. Yang, J.-M. Lee, and M. J. Kim, "Decomposition-based multi-objective firefly algorithm for rfid network planning with uncertainty," *Applied Soft Computing*, vol. 55, pp. 549–564, 2017.

[100] S. Zhu and M. Fukushima, "Worst-case conditional value-at-risk with application to robust portfolio management," *Operations research*, vol. 57, no. 5, pp. 1155–1168, 2009.

[101] S. Zhu, D. Li, and S. Wang, "Robust portfolio selection under downside risk measures," *Quantitative Finance*, vol. 9, no. 7, pp. 869–885, 2009.

[102] W. T. Ziemba, W. Ziemba, J. M. Mulvey, J. Mulvey, H. Moffatt *et al.*, *Worldwide asset and liability modeling.* Cambridge University Press, 1998, vol. 10.

[103] S. Zymler, D. Kuhn, and B. Rustem, "Distributionally robust joint chance constraints with second-order moment information," *Mathematical Programming*, vol. 137, no. 1-2, pp. 167–198, 2013.

[104] S. Zymler, B. Rustem, and D. Kuhn, "Robust portfolio optimization with derivative insurance guarantees," *European Journal of Operational Research*, vol. 210, no. 2, pp. 410–424, 2011.

Every reasonable effort has been made to acknowledge the owners of copyright material. I would be pleased to hear from any copyright owner who has been omitted or incorrectly acknowledged.

19 Oct 2020

Journal of Industrial and Management Optimization

Dear Editor-in-Chief:

I am completing a thesis at Curtin University entitled "Robust and Robust and Multi-objective Portfolio Selection." I would like your permission to reprint in my dissertation excerpts from the following:

> JIANG, Lin; WANG, Song. Robust multi-period and multi-objective portfolio selection. Journal of Industrial & Management Optimization, 2019, 13.5: 0.
>
> The excerpts to be reproduced as a chapter in my thesis. The requested permission covers future revisions and editions of my dissertation and to the prospective publication of my dissertation as an electronic dissertation available through the Library at Curtin University. These rights will in no way limit republication of the material(s) in any other form by you or others authorized by you. Your signing will verify that you own the copyright to the above material(s).

If this meets with your approval, please sign this letter below and return it to me through email. Thank you very much for your attention to this matter.

Sincerely,

Lin Jiang

Enclosure

PERMISSION GRANTED FOR THE USE REQUESTED ABOVE:

[redacted]

EIC of Industrial and Management Optimization

[Type name of addressee below signature line]

Date: 21 October, 2020

Gmail

lin Jiang <lydiajianglin@gmail.com>

your thesis examinations

Lin Jiang <lydiajianglin@gmail.com> 2020年10月19日 上午10:57
收件人：k.l.teo@curtin.edu.au, Liwei Ning <editorial@aimsciences.org>, Susan Cummins <journal@aimsciences.org>
抄送：Song Wang <Song.Wang@curtin.edu.au>

Dear all

I have a publication "Robust multi-period and multi-objective portfolio selection" with my supervisor Prof. Song Wang in JIMO. I have used part of this publication to finish my PhD thesis. According to our university policy, I need a permission form from the publisher to allow me to do this. Can you please do me a favor and give me evidence to allow me to do this?

Many thanks

Lin

Song Wang <Song.Wang@curtin.edu.au> 于2020年10月19日周一 上午10:29写道：
[引用文字已隐藏]

Gmail

lin Jiang <lydiajianglin@gmail.com>

your thesis examinations

Liwei Ning <editorial@aimsciences.org> 2020年10月22日 下午10:24
收件人：lydiajianglin@gmail.com
抄送：Susan Cummins <journal@aimsciences.org>, Liwei Ning <editorial@aimsciences.org>

Dear Lin Jiang,

You have the permission provided the proper acknowledgement is made of the original publication. Thanks.

Best regards,

Liwei Ning

Liwei Ning

Editorial Manager

American Institute of Mathematical Sciences

Tel/Fax: 417-351-3204

Email: editorial@aimsciences.org

[引用文字已隐藏]