Linköping Studies in Science and Technology. Theses
No. 1717

Mean-Variance Portfolio Optimization: Eigendecomposition-Based Methods

Fred Mayambala

INSTITUTE OF TECHNOLOGY
LINKÖPINGS UNIVERSITET

Department of Mathematics
Linköping University, SE-581 83 Linköping, Sweden
Linköping 2015

Linköping Studies in Science and Technology. Theses
No. 1717

Mean-Variance Portfolio Optimization: Eigendecomposition-Based Methods

Fred Mayambala

fred.mayambala@liu.se

www.mai.liu.se

Division of Optimization

Department of Mathematics

Linköping University

SE-581 83 Linköping

Sweden

ISBN 978-91-7519-038-9

ISSN 0345-7524

Copyright © 2015 Fred Mayambala

Printed by LiU-Tryck, Linköping, Sweden 2015

To the International Science Programme (ISP)

Abstract

Modern portfolio theory is about determining how to distribute capital among available securities such that, for a given level of risk, the expected return is maximized, or for a given level of return, the associated risk is minimized. In the pioneering work of Markowitz in 1952, variance was used as a measure of risk, which gave rise to the well-known mean-variance portfolio optimization model. Although other mean-risk models have been proposed in the literature, the mean-variance model continues to be the backbone of modern portfolio theory and it is still commonly applied. The scope of this thesis is a solution technique for the mean-variance model in which eigendecomposition of the covariance matrix is performed.

The first part of the thesis is a review of the mean-risk models that have been suggested in the literature. For each of them, the properties of the model are discussed and the solution methods are presented, as well as some insight into possible areas of future research.

The second part of the thesis is two research papers. In the first of these, a solution technique for solving the mean-variance problem is proposed. This technique involves making an eigendecomposition of the covariance matrix and solving an approximate problem that includes only relatively few eigenvalues and corresponding eigenvectors. The method gives strong bounds on the exact solution in a reasonable amount of computing time, and can thus be used to solve large-scale mean-variance problems.

The second paper studies the mean-variance model with cardinality constraints, that is, with a restricted number of securities included in the portfolio, and the solution technique from the first paper is extended to solve such problems. Near-optimal solutions to large-scale cardinality constrained mean-variance portfolio optimization problems are obtained within a reasonable amount of computing time, compared to the time required by a commercial general-purpose solver.

Populärvetenskaplig sammanfattning

För den som har kapital att investera kan det vara svårt att avgöra vilka investeringar som är mest fördelaktiga. Till stöd för beslutet kan matematiska modeller användas och denna avhandling handlar om hur man kan beräkna lösningar till sådana modeller. De investeringsalternativ som betraktas är finansiella instrument som är föremål för daglig handel, som aktier och obligationer.

En investerare placerar kapital i finansiella instrument eftersom de förväntas ge en god avkastning över tiden. Samtidigt är sådana placeringar alltid förknippade med risktagande. Förväntad avkastning och risk varierar kraftigt mellan olika instrument. Till exempel ger placeringar i statsobligationer typiskt mycket låg avkastning till mycket låg risk, medan placeringar i aktier i nystartade bolag som utvecklar nya läkemedel kan ge mycket hög avkastning samtidigt som risken är mycket hög.

Att investera kan ses som en avvägning mellan den förväntade avkastningen och den risk som investeringen innebär, och typiskt är hög förväntad avkastning också associerade med en hög risk, vilken kan leda till stora förluster. En rationell investerare vill undvika alltför stora risker, men för att investeringen ska bli rimligt lönsam måste en viss risk accepteras.

För att minska den totala risken sprider en investerare normalt sitt kapital på en portfölj av finansiella instrument. Dock är vanligen avkastningarna för instrumenten i en portfölj inte oberoende av varandra, utan samvarierar. Till exempel kan alla bolag inom en och samma bransch förväntas ha likartade beroenden av den ekonomiska konjunkturen. Denna omständighet försvårar avsevärt problemet att sätta samman en portfölj. Instrumenten och de kapital som investeras i var och en av dem väljs så att både den samlade avkastningen och den samlade risken för portföljen blir acceptabel utifrån investerarens preferenser.

Matematiska modeller som kan användas för att finna en portfölj av investeringar som är optimal med avseende på den önskade avvägningen mellan förväntad avkastning och risk med de investeringsalternativ som finns tillgängliga på marknaden är typiskt beräkningskrävande, samtidigt som man på kort tid vill kunna ta fram flera olika förslag på portföljer. I denna avhandling presenteras en ny typ av beräkningsmetoder som är bra på att ta fram optimala portföljer på kort tid.

Acknowledgments

Now unto him that is able to do exceeding abundantly above all that we ask or think, according to the power that worketh in us, unto him be glory in the church by Christ Jesus throughout all ages, world without end. Amen. (Ephesians 3:20-21). I will always glorify your holly name through your son, Jesus Christ. All that is possible for me, is for the glory of the most high.

I have special thanks for my supervisor Torbjörn Larsson. Never in my life had I ever interacted with an extremely intelligent person like you. Your perfect combination of hardwork and intelligence make you a special person to work with. Besides being my academic supervisor, on very many occasions you treated me like a parent, and that will always keep in my heart. I honestly just can't thank you enough. I just hope that our collaboration continues.

You introduced me to my co-supervisor, Elina Rönnerberg, who has been very helpful to me. For most of the times that I have met Elina, something has been completed. You have always given this work a good direction. I am very grateful to you, Elina.

I want also to thank my other co-supervisor, Juma Kasozi, whom I mostly worked with when I went back to Uganda. Its now ten years since I first became your student. You have always treated me in a special way and that is why, I think, we are still working together up to now.

If it were not for Bent-Ove Turesson, I would have probably ended up in another univeristy in Sweden. Thank you so much for giving me this great chance to come to Linköping University, and also for being kind to me, whenever I came to you. I have been well looked after by Björn Textorius. I thank you so much Björn for making me feel at home and always helping whenever I needed your assistance. Theresa, it has always been a pleasure meeting you. I want also to thank all the members of the department that I have interacted with in various ways, and all my fellow students at the Department of Mathematics for being such a wonderful group of people.

I am grateful to my family for always accepting the pain of missing me for long periods. I thank you mum Rose Kyomugisha for being a hero in my life. I also thank you grand mam Manjeri: you have been the father figure in my life. I also thank my girl, Becky; you always encouraged me when the going went tough.

All this work would never have been possible if it were not for the financial support from the International Science Programme (ISP). You found a toddler that was trying to walk, and you have not only taught me how to walk, but you have also showed me the path to walk from. During my study, I have mainly interacted with Pravina and Leif, from Sweden. You have been very cooperative and helpful to me. I thank you so much and kindly request you to convey my appreciation to ISP. I also thank the ISP coordinators in Uganda, John Mango and Juma Kasozi, who head the Eastern African Universities Mathematics Programme.

Linköping, June 9, 2015

Fred Mayambala

Contents

1	Introduction	1
1.1	Background	1
1.2	Outline	2
1.2.1	Outline of Part I	2
1.2.2	Outline of Part II	2
I	Mean-Risk Portfolio Optimization	5
2	Mean-Risk Models	7
2.1	Mean	7
2.2	Risk	7
2.2.1	Risk Measurement	8
2.2.2	Risk Aversion	9
2.3	Mean-Risk Models in Portfolio Optimization	10
3	Review	11
3.1	Mean-Variance Model	11
3.1.1	Transaction Costs in the Mean-Variance Model	14
3.1.2	Cardinality Constrained Optimization Models	16
3.1.3	Robust Optimization in the Mean-Variance Model	20
3.1.4	Multi-Period Mean-Variance Optimization	24
3.2	Mean Absolute Deviation Model	28
3.2.1	MAD under Transaction Costs	29
3.2.2	A Robust MAD Model	30
3.3	Value-at-Risk Model	31
3.4	Conditional Value-at-Risk Model	33

3.4.1	Robust Optimization Using CVaR	36
3.5	Mean Absolute Semi Deviation Model	37
3.5.1	MASD with Real Features	37
4	Concluding Remarks	39
	Bibliography	41
II	Research Papers	53
A	Paper A	55
1	Introduction	58
2	Eigendecomposition of the Mean-Variance Model	59
2.1	Approximation of the Mean-Variance Model	60
3	An Improved Approximation Strategy	62
3.1	A Linearized Error Term	63
3.2	Cardinality of the Solution	66
4	Numerical Illustrations	67
4.1	Upper and Lower Bounds	68
4.2	Deviation in Solution	69
4.3	Efficient Frontier	71
4.4	Cardinality of the Solution	72
5	A Proposed Transformation	74
6	Conclusion and Further Research	76
	References	79
B	Paper B	81
1	Introduction	84
2	Theory and Solution Principle	84
3	Numerical Results and Conclusion	87
	References	91

1

Introduction

The mean-risk hypothesis which was proposed by Markowitz in 1952 has had a profound effect in both the research world and the practical financial world. A lot of research has been rooted from this great hypothesis and a number of mean-risk models have been developed. The solution methods employed to solve these models are quite diverse and the interesting thing is that some of them still require better solution techniques in terms of computational time, ease to handle large-scale problems and ease to implement or use. In the thesis, we use a method based on eigendecomposition to solve a mean-variance model with different sets of constraints.

1.1 Background

The break-through of modern portfolio theory was realised with a seminal paper by Markowitz [100]. As opposed to the thinkings at that time, Markowitz argued that an investor who aimed at getting higher returns also needed to mind about the risks involved, leading to the “mean-risk” hypothesis. Markowitz’s model was the first mean-risk model in portfolio optimization. As a measure of risk, Markowitz decided to use standard deviation or variance of the portfolio returns. Soon questions rose on whether variance was an appropriate measure of risk and Markowitz suggested semi-variance as an alternative risk measure [101]. Nevertheless mean-variance models still gained popularity amongst researchers, even to date. Later Markowitz’s model was modified to include more realistic features like transaction costs (see [112], [113], [93], [8]), cardinality constraints (see [14], [127], [49]), and multi-period optimization (see [87], [81], [84]).

A portfolio optimization model to capture the preference of the investor was suggested by Tobin [138]. The preference of the investor can be captured in a function called a utility function. By maximizing expected utility, an optimal portfolio for risk averse investors could be determined. Tobin’s work soon gained a lot of interest from researchers (see [107], [56], [44] for multi-period utility maximization models, and [35], [50], [108] for

transaction costs). There was a general perception that mean-variance optimization was in fact a special case of expected utility maximization with a quadratic utility function, a fact which some authors objected (see [83], [80]).

However, due to the computational burden of the Markowitz model, linear models were sought and this saw the birth of a mean absolute deviation model ([68], [78]) in which mean absolute deviation was used as a risk measure. The mean absolute deviation model also gained a lot of popularity in the 1990s (see for example [71], [69], [105]).

A very popular measure of risk in the 1990s called value-at-risk also found usage in portfolio optimization (see [54], [128]). Value-at-risk had been accepted by the Basle committee as a good measure of risk for financial institutions. However, value-at-risk has some undesirable properties [5] and is therefore not a very good choice in portfolio optimization. The undesirable properties of value-at-risk led to the birth of a new risk measure, called conditional value-at-risk, with better properties than value-at-risk. After its introduction by Rockafeller and Uryasev [121], conditional value at risk gained a lot of interest from researchers, see [79], [119], [148].

Other risk measures have been developed, see for example [132], [17].

1.2 Outline

The thesis consists of two parts, outlined as follows.

1.2.1 Outline of Part I

This part consists of two chapters.

Chapter 2 provides background material for the mean-risk models. The concepts of mean and risk are explained. The chapter ends with a general mean-risk model.

Chapter 3 is a review of the mean-risk models in portfolio optimization. The risk models considered are variance, mean absolute deviation, value-at-risk, conditional value-at-risk and the mean-mean absolute semi deviation. For all these models, we look at the extensions of the model to include real features like cardinality constraints, transaction costs, robust models and multi-period models. The survey covers papers from 1952 up to date.

1.2.2 Outline of Part II

This part consists of two papers. Below is a brief outline for each of the papers

Paper A: Eigendecomposition of the Mean-Variance Portfolio Optimization Model

This paper provides a new insight into the mean-variance portfolio optimization problem, based on performing eigendecomposition of the covariance matrix. We show that when only some of the eigenvalues and eigenvectors are used, the resulting mean-variance problem is an approximation of the original one. The approximate mean-variance model obtained actually gives good lower and upper bounds when tested on real world stock

market data from NYSE. We also propose techniques to further improve the obtained bounds. One of the techniques is to include a linearized error term.

We also propose an ad hoc linear transformation of the mean-variance problem, which in practice significantly strengthens the bounds obtained from the approximate mean-variance problem

Paper B: Tight Upper Bounds on the Cardinality Constrained Mean-Variance Portfolio Optimization Problem Using Truncated Eigendecomposition

This paper introduces a core problem based method for obtaining upper bounds to the mean-variance portfolio optimization problem with cardinality and bound constraints. Like paper A, this paper also involves performing eigendecomposition on the covariance matrix and then using only few of the eigenvalues and eigenvectors to obtain an approximation of the original problem.

The core problem is formed based on the approximate mean-variance problem and used to obtain upper bounds on the cardinality constrained mean-variance problem. The method is tested on real world data from the NYSE market and the computation time is observed to be very promising.

Part I

Mean-Risk Portfolio Optimization

2

Mean-Risk Models

This chapter focuses on the background material in portfolio optimization. The concept of risk and mean are handled, before ending the chapter with a general form of a mean-risk model.

2.1 Mean

Throughout the thesis, unless stated otherwise, we shall consider a portfolio of n securities in which an investor invests a fraction x_i , $i = 1, 2, \dots, n$, of the available funds into the i^{th} security. The rate of return on the i^{th} investment shall be a random variable r_i with $E(r_i) = \mu_i$, $i = 1, 2, \dots, n$. Letting $\mathbf{x} = (x_1, x_2, \dots, x_n)^T$ and $\mathbf{r} = (r_1, r_2, \dots, r_n)^T$, the return on the portfolio is then

$$R = \mathbf{x}^T \mathbf{r}. \quad (2.1)$$

Letting $\boldsymbol{\mu} = (\mu_1, \mu_2, \dots, \mu_n)^T$, the expected return, also called mean, of the portfolio is given by

$$E(R) = E\left(\sum_{i=1}^n x_i r_i\right) = \sum_{i=1}^n x_i E(r_i) = \sum_{i=1}^n x_i \mu_i = \mathbf{x}^T \boldsymbol{\mu}.$$

2.2 Risk

Risk is any event or action that may adversely affect an organization's or individual's ability to achieve its objectives and execute its strategies [102]. Risks can be broadly grouped into two categories.

Business risks: These are risks that organizations take on willingly in order to add value to the organization. For example strategic risks.

Financial risks: These are risks that arise due to possible losses which are entirely market driven. Financial risks can be subdivided into market risk, credit risk and operational risk.

For a detailed study on types of financial risks, see [62], [63] and [102]. In portfolio optimization, the most common types of risks studied are the financial risks, and they are the main issue of concern in this work.

2.2.1 Risk Measurement

Let $\mathcal{G}$ denote the set of all possible risks. Then a risk measure $\mathcal{R}$ is a mapping from $\mathcal{G}$ into $\mathfrak{R}$. Common examples of risk measures include

- **dispersion risk measures** which measure the dispersion of the random variables (gains or losses) from a parameter, for example standard deviation [100] or mean absolute deviation [78].
- **downside risk measures** which are associated with the worst outcome being below some set target and its probability, for example value-at-risk, conditional value-at-risk, lower semi-variance, and
- **sensitivities** which measure the sensitivity of change of value of securities when small changes are made in the underlying parameters, for example delta, vega, theta and others used in measuring sensitivity of derivative prices.

A detailed coverage of risk measures is given in the survey [6].

Coherent Risk Measures

Definition 2.1 (Artzner et al [5]). A risk measure $\mathcal{R}$ is said to be coherent if it satisfies axioms **P1-P4** below.

P1: Translation invariance: For all $\mathbb{X} \in \mathcal{G}$ and all $\alpha \in \mathfrak{R}$

$$\mathcal{R}(\mathbb{X} + \alpha) = \mathcal{R}(\mathbb{X}) - \alpha,$$

which means that if the initial amount to be invested into an asset is made larger, then risk will reduce.

P2: Sub-additivity: For all $\mathbb{X}_1$ and $\mathbb{X}_2$ in $\mathcal{G}$,

$$\mathcal{R}(\mathbb{X}_1 + \mathbb{X}_2) \leq \mathcal{R}(\mathbb{X}_1) + \mathcal{R}(\mathbb{X}_2),$$

which means that diversification will lead to a reduction in risk.

P3: Positive homogeneity: For all $\mathbb{X} \in \mathcal{G}$ and $\beta \geq 0$,

$$\mathcal{R}(\beta \mathbb{X}) = \beta \mathcal{R}(\mathbb{X}),$$

which means that increasing the portfolio leads to a proportionally increased risk.

P4: Monotonicity: For all $\mathbb{X}_1, \mathbb{X}_2 \in \mathcal{G}$, with $\mathbb{X}_1 \leq \mathbb{X}_2$,

$$\mathcal{R}(\mathbb{X}_2) \leq \mathcal{R}(\mathbb{X}_1),$$

meaning that a risk measure should be a monotonically increasing function with increasing risk.

Some risk measures are coherent while others are not, as we shall see in the later sections. For discrete time coherent risk measures, see [26], [27], for coherent risk measures on general probability spaces, see [36], and [38] for coherent risk allocation.

2.2.2 Risk Aversion

An agent is said to be risk-averse if at any wealth level, he/she dislikes every lottery with an expected pay-off of zero [41]. The agent's preference towards risk can be modelled using a utility function. For an agent to be risk averse, the utility function should be concave. For more details on utility functions, an interested reader is referred to [140]. If u is a utility function, then at any wealth level w and zero mean lottery ε , an agent is said to be risk averse if

$$E(u(w + \varepsilon)) < u(w).$$

One way of determining the degree of risk is to determine the risk premium P , associated to that risk using the expression

$$E(u(w + \varepsilon)) = u(w - P). \quad (2.2)$$

By using the second order Taylor expansion on the left hand side of (2.2) and the first order Taylor series expansion on the right hand side, the risk premium P would satisfy

$$P \approx \frac{1}{2} E(\varepsilon^2) A(w), \quad (2.3)$$

where $E(\varepsilon^2)$ is the variance of the outcome of the lottery and

$$A(w) = -\frac{u''(w)}{u'(w)}. \quad (2.4)$$

Equation (2.4) is called the Arrow-Pratt measure of absolute risk aversion [117]. Based on (2.4), an agent is risk averse if $A(w) > 0$, risk neutral if $A(w) = 0$ and risk loving if $A(w) < 0$. Another measure of risk aversion is the relative risk aversion $R(w)$ given by

$$R(w) = -w \frac{u''(w)}{u'(w)} = wA(w).$$

Since different agents have different preferences, it is common to assign different utility functions to different agents based on their preferences. Common utility functions in literature include the following.

- Quadratic utility function: $u(w) = w - \frac{b}{2}w^2$, $b > 0$.
- Constant absolute risk-aversion utility function; $u(w) = -\frac{e^{-aw}}{w}$.
- Power utility functions;

$$u(w) = \begin{cases} \frac{w^{1-\lambda}}{1-\lambda}, & \lambda \geq 0, \lambda \neq 1 \\ \ln(w), & \lambda = 1. \end{cases}$$

Kroll et al. [80] establish the connection between different utility functions and variance.

2.3 Mean-Risk Models in Portfolio Optimization

According to [100], a portfolio is said to belong to an *efficient frontier* if for a given level of expected return, it has minimum risk, and for a given level of risk, it has maximum expected return.

Suppose that we choose a minimum level of expected return of the portfolio as μ_P , a maximum level of risk as σ_P^2 , and $\mathcal{S}$ as the set of all possible portfolios. Then the mean-risk model takes on any of the three forms.

$$\begin{array}{llll}
 \min_{\mathbf{x}} & \mathcal{R}(\mathbf{R}) & \max_{\mathbf{x}} & \mathbf{x}^T \boldsymbol{\mu} \\
 \text{s.t.} & \mathbf{x}^T \boldsymbol{\mu} \geq \mu_P & \text{s.t.} & \mathcal{R}(\mathbf{R}) \leq \sigma_P^2 \\
 & \mathbf{x} \in \mathcal{S} & & \mathbf{x} \in \mathcal{S}
 \end{array} \quad (2.5) \quad (2.6) \quad (2.7)$$

The parameter λ in (2.7) is called a risk averseness parameter. All the mean-risk models take on the general form (2.5), (2.6) or (2.7). The equivalence between the optimal solutions of any of (2.5), (2.6) or (2.7), can be established by fixing the parameters μ_P , σ_P^2 , or λ , for the three models respectively.

3

Review

In this chapter we provide a review of the most common mean-risk models in the literature from 1952 up to present. These models basically take on the forms (2.5), (2.6) or (2.7), with different measures of risk, $\mathcal{R}$. In Section 3.1, we consider a model in which $\mathcal{R}$ is the variance. The case where $\mathcal{R}$ is the mean absolute deviation is considered in Section 3.2, value-at-risk is considered in 3.3, conditional value-at-risk is considered in 3.4, and mean absolute semi deviation is considered in Section 3.5.

3.1 Mean-Variance Model

By definition, the variance of a random variable R is

$$\text{Var}(R) = E((R - E(R))^2).$$

But since an expression for R is given in (2.1), we can use the property of the variance of the sum of random variables, so that the variance of the total return on portfolio is given by

$$\text{Var}(R) = \sum_{i=1}^n \sum_{j=1}^n x_i x_j \sigma_{ij} = \mathbf{x}^T \mathbf{\Sigma} \mathbf{x},$$

where σ_{ij} is the covariance between the i^{th} and j^{th} asset returns, and $\mathbf{\Sigma}$ is the $n \times n$ symmetric positive semi-definite ($\mathbf{\Sigma} \succeq 0$) matrix of covariances. Variance is not a coherent risk measure but rather a deviation risk measure [120].

Let us define a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, where Ω is a sample space, $\mathcal{F}$ is a measurable space on Ω and $\mathbb{P}$ is a probability measure on $\mathcal{F}$. Then $\mathcal{L}^2(\Omega)$ is an L_2 space of random portfolio returns. Let $R, Y \in \mathcal{L}^2(\Omega)$, where Y represents constant random variables on $\mathcal{L}^2(\Omega)$.

Theorem 3.1

Let f be a functional defined by

$$f(R, Y) = E((R - Y)^2),$$

then f attains minimum at $Y = E(R)$ and the minimum is the variance of R .

Proof: First note that $f(R, Y) = E((R - Y)^2) = E(R^2) - 2E(R)Y + Y^2$. Therefore, differentiating f and equating to zero gives the result $Y = E(R)$ and the definition of variance. $\square$

Thus Theorem 3.1 shows that the variance of a random variable is the minimum distance (in mean square sense) of a random variable from its expected value. A portfolio $\mathbf{x}$ with minimum variance is therefore one which gives the least distance (in mean square sense) between the portfolio return R , and $E(R)$.

There exists a variety of portfolio combinations, each of which having its own portfolio return. The question which arises is: among all possible portfolio combinations, which one has the minimum distance, in least square sense, between the return on the portfolio and the portfolio expected return? This is an optimization problem

$$\min_{\mathbf{x}} f(R, E(R)) \quad \implies \quad \min_{\mathbf{x}} \mathbf{x}^T \Sigma \mathbf{x}.$$

However, with the assumption that $\Sigma \succeq 0$, an optimal solution to such a problem is $\mathbf{x}^* = 0$, which translates to zero portfolio returns. The realistic investor's problem is to set a required positive level of portfolio returns, say to μ_P , and the investor's portfolio optimization problem (assuming no other constraints are present) thus becomes

$$\begin{aligned} \min_{\mathbf{x}} \quad & \mathbf{x}^T \Sigma \mathbf{x} \\ \text{s.t.} \quad & \boldsymbol{\mu}^T \mathbf{x} = \mu_P. \end{aligned} \tag{3.1}$$

Using Lagrange multipliers (see [111] for details on using Lagrange multipliers), the optimal solution to problem (3.1) is

$$\mathbf{x}^* = \frac{\Sigma^{-1} \boldsymbol{\mu} \mu_P}{\boldsymbol{\mu}^T \Sigma^{-1} \boldsymbol{\mu}},$$

if Σ^{-1} exists. Note however that changing the second equation in (3.1) to $\boldsymbol{\mu}^T \mathbf{x} \geq \mu_P$ does not change the optimal solution because the optimal solution will still be attained with an equality in the constraint. This observation follows from Theorem 3.1. This is because a value of $\boldsymbol{\mu}^T \mathbf{x}$ greater than μ_P will give a larger variance.

The solution of (3.1) puts no bound on the amount of capital available to an investor. To cater for a limit on the capital, a constraint $\mathbf{e}^T \mathbf{x} = 1$, where $\mathbf{e}$ is a vector of ones, is added to problem (3.1). This means that each x_i , $i = 1, 2, \dots, n$ represents a fraction of capital invested in the i^{th} asset. Thus the problem becomes

$$\begin{aligned} \min_{\mathbf{x}} \quad & \mathbf{x}^T \Sigma \mathbf{x} \\ \text{s.t.} \quad & \boldsymbol{\mu}^T \mathbf{x} = \mu_P \\ & \mathbf{e}^T \mathbf{x} = 1. \end{aligned} \tag{3.2}$$

Problem (3.2) is called the *mean-variance boundary problem* [54]. Using Lagrange multipliers, the optimal solution of (3.2), which is given in [103], is

$$\mathbf{x}^* = \frac{A(\mu_P - A)}{D} \frac{\boldsymbol{\Sigma}^{-1} \boldsymbol{\mu}}{A} + \frac{C(B - A\mu_P)}{D} \frac{\boldsymbol{\Sigma}^{-1} \mathbf{e}}{C}$$

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

The addition of a lower bound on the asset holdings $\mathbf{x}$, in the model (3.2) eliminates the possibility of using Lagrange multipliers to solve the problem. The view that investors' choices of investment are influenced by both expected portfolio returns and associated risk was the basis of the Markowitz model and leads to the following models (see [100] and [101]).

$$\begin{array}{llll} \min_{\mathbf{x}} & \mathbf{x}^T \boldsymbol{\Sigma} \mathbf{x} & \max_{\mathbf{x}} & \mathbf{x}^T \boldsymbol{\mu} \\ \text{s.t.} & \mathbf{x}^T \boldsymbol{\mu} \geq \mu_P & \text{s.t.} & \mathbf{x}^T \boldsymbol{\Sigma} \mathbf{x} \leq \sigma_P^2 \\ & \mathbf{e}^T \mathbf{x} = 1 & & \mathbf{e}^T \mathbf{x} = 1 \\ & \mathbf{x} \geq 0 & & \mathbf{x} \geq 0 \end{array} \quad (3.3) \quad (3.4) \quad (3.5)$$

Here μ_P is the lowest accepted target for portfolio returns, σ_P^2 is the maximum allowed variance for the portfolio returns and $\lambda > 0$ is regarded as a risk averseness parameter. For a fixed $\lambda > 0$, if $\mathbf{x}(\lambda)$ solves (3.5), then it also solves (3.3) if μ_P is chosen as $\boldsymbol{\mu}^T \mathbf{x}(\lambda)$, and solves (3.4) if σ_P^2 is chosen as $\mathbf{x}(\lambda)^T \boldsymbol{\Sigma} \mathbf{x}(\lambda)$ [101]. The term *efficient frontier* is used to refer to a set of portfolios with the property that for every such portfolio, for a given expected return level, it has minimum variance and for a given variance, it has maximum expected returns. With the emergence of a variety of softwares to solve quadratic programming problems, models (3.3), (3.4) and (3.5), can readily be solved at least if the problem is not too large.

However, with a very dense $\boldsymbol{\Sigma}$ and a large number of assets, problems (3.3), (3.4) and (3.5), can take a large amount of time to solve. This problem was noticed in the 1980s and the initial works by Perold [113] paved way for future work. The use of a parametric quadratic programming approach on large non-factorable covariance matrices by Perold, was deemed ineffective by Kawada and Konno [64], who proposed to decompose the variance into seperable functions. After obtaining a starting point, the method involved using a steepest descent algorithm and a variable metric algorithm to obtain an optimal solution. The incorporation of a third moment of R , the skewness, as seen in [124] reduces the number of variables considerably and the resulting model can be used to solve large scale problems. An active set method is proposed [135] to solve large scale versions of (3.3). However, the most effective method as seen from the survey by [136] is a method proposed in [118] and [58] which uses a parametric method to compute a solution of the problem (3.5). This method obtains a single solution for the entire efficient frontier. The surprise is that it is never used in practice to compute the efficient frontier. This could possibly be due to the fact that the method is too specialized for only efficient frontiers and, like the critical line method which was proposed in [101], it has been overshadowed by the more interesting interior point methods which exist in many of the state-of-the-art softwares.

There is still more need to devise more efficient quadratic programming algorithms to

solve large scale versions of the mean-variance problems, to match the growing sizes of financial markets in the world.

Other constraints can be added to the problem, for example $l_i \leq x_i \leq u_i$ $i = 1, \dots, n$, which is called a transaction level constraint and limits the amount to be invested in each asset. For the rest of the work, unless stated otherwise, we use the set $\mathcal{S}$ to denote the set of all possible investments. The constraint $\mathbf{e}^T \mathbf{x} = 1$ shall always be assumed to be part of the set $\mathcal{S}$.

3.1.1 Transaction Costs in the Mean-Variance Model

Transaction costs are basically any costs incurred when buying or selling securities, for example brokerage fees, taxes, bid-ask spreads, and so on. We denote transaction costs incurred on trading in the i^{th} security by $\phi_i(x_i)$ and the total transaction costs on the portfolio by $\phi(\mathbf{x})$. Transaction costs that can be incurred while transacting are of two types.

- (i) Fixed transaction costs: these are incurred by any investor who chooses to transact business in an asset and are independent of the amount traded. We shall denote such a cost by F_i . We shall denote the fixed transaction cost on selling and buying by F_i^S and F_i^B respectively.
- (ii) Variable transaction costs: these costs depend on the amount traded in an asset. Let f_i^S and f_i^B denote the variable transaction cost functions associated with selling and buying, respectively, of the i^{th} asset.

Let $[b_{ij}, b_{ij+1}]$, $j = 0, 1, \dots, B_i$, be intervals with the same variable transaction cost function where B_i is the total number of such intervals for the i^{th} asset. Let M be the total amount of capital available to the investor so that Mx_i is the amount of capital invested in the i^{th} asset. Then the total transaction cost ϕ_i , including both fixed and variable transaction costs, for the i^{th} asset is, in general, given by

$$\phi_i(x_i) = \begin{cases} 0 & x_i = 0 \\ F_i^B + f_i^S(x_i), & \text{if } x_i > 0 \text{ and } b_{ij} < Mx_i \leq b_{ij+1}, \quad j = 0, 1, \dots, B_i \\ F_i^S - f_i^S(x_i), & \text{if } x_i < 0 \text{ and } \bar{b}_{ij} < Mx_i \leq \bar{b}_{ij+1}, \quad j = 0, 1, \dots, B_i. \end{cases} \quad (3.6)$$

The bar on the intervals for $x_i < 0$ are used to distinguish them from those for $x_i > 0$. Clearly the function $\phi_i(x_i)$ is in general, a discontinuous, non-linear and non-differentiable function. These properties make it very difficult to solve optimization problems with general transaction costs (3.6). The total transaction cost for the portfolio is thus the separable function

$$\phi(\mathbf{x}) = \sum_{i=1}^n \phi_i(x_i). \quad (3.7)$$

There are different ways of embedding transaction costs into the portfolio optimization problem (3.3), (3.4) and (3.5), some of which are

$$\begin{aligned}
\max_{\mathbf{x}} \quad & \boldsymbol{\mu}^T \mathbf{x} - \phi(\mathbf{x}) \\
\text{s.t.} \quad & \mathbf{x}^T \boldsymbol{\Sigma} \mathbf{x} \leq \sigma_P^2 \\
& \mathbf{x} \in \mathcal{S}
\end{aligned} \quad (3.8)$$

$$\begin{aligned}
\max_{\mathbf{x}} \quad & \boldsymbol{\mu}^T \mathbf{x} - \lambda \mathbf{x}^T \boldsymbol{\Sigma} \mathbf{x} - \phi(\mathbf{x}) \\
& \mathbf{x} \in \mathcal{S}
\end{aligned} \quad (3.9)$$

$$\begin{aligned}
\min_{\mathbf{x}} \quad & \phi(\mathbf{x}) \\
\text{s.t.} \quad & \mathbf{x}^T \boldsymbol{\Sigma} \mathbf{x} \leq \sigma_P^2 \\
& \boldsymbol{\mu}^T \mathbf{x} \geq \mu_P \\
& \mathbf{x} \in \mathcal{S}.
\end{aligned} \quad (3.10)$$

It should be noted that all the models (3.8), (3.9) and (3.10) are non-convex problems. In the literature, most of the work involving transaction costs has been aimed at putting simplifying assumptions on (3.6) to end up with problems that are much easier to solve. Below are some of these assumptions.

Patel and Subrahmanyam [112] consider model (3.9) under fixed transaction costs only and with another simplifying assumption in their model, that all securities included in the portfolio will carry equal fixed transaction cost α . This is achieved by setting

$$\phi(\mathbf{x}) = \sum_{i=1}^n \phi_i(x_i) = \alpha \sum_{i=1}^n y_i \quad (3.11)$$

$$\text{with } y_i = \begin{cases} 1, & \text{if } x_i \neq 0, \quad i = 1, 2, \dots, n \\ 0, & \text{otherwise.} \end{cases}$$

The set $\mathcal{S}$ is also assumed to contain the constraint $\mathbf{e}^T \mathbf{x} = 1$ only. Clearly embedding (3.11) into (3.9) leads to a mixed-integer quadratic programming (MIQP) problem. However, assuming equal asset correlation coefficients ρ_{ij} , ($\rho_{ij} = \sigma_{ij}/\sigma_i\sigma_j$), [112] devise a simpler algorithm that avoids the direct solution of the MIQP problem.

Perold [113] uses transaction cost function (3.7) in which each $\phi_i(x_i)$ is a concave and piecewise linear function. The resulting problem is solved using a parametric algorithm.

Xue et al. [146] use the transaction cost (3.7) in which each $\phi_i(x_i)$ is a non-decreasing concave function. With the assumption that transaction costs are smooth enough, the resulting cost function is a difference of two convex functions. The resulting problem is solved using a branch-and-bound algorithm.

Lobo et al. [93] consider transaction cost (3.7) in which each $\phi_i(x_i)$ is defined as

$$\phi_i(x_i) = \begin{cases} 0 & x_i = 0 \\ F_i^B + \alpha_i^B x_i, & x_i > 0 \\ F_i^S - \alpha_i^S x_i, & x_i < 0, \end{cases} \quad (3.12)$$

where $\alpha_i^B x_i$ and $\alpha_i^S x_i$ are proportional transaction costs associated with buying and selling respectively, the i^{th} asset. The resulting non-convex problem is solved using a heuristic method.

Bertsimas and Shioda [8] use transaction cost (3.7) in which each $\phi_i(x_i)$ is a quadratic function of the current and new portfolio portions. Together with cardinality constraints, the resulting MIQP problem is solved using a branch-and-bound algorithm.

Assuming that $F_i^B = F_i^S = 0$, so that each $f_i(x_i)$ is a piece-wise linear function, then the non-differentiable function $f_i(x_i)$ can be approximated by a smooth function to give a convex problem. Potapchik et al. [116] use convex spline functions to approximate $f_i(x_i)$

and then solve the resulting problem using a combination of both interior point and active set methods.

Other forms of (3.7) are considered in [89], [82], [88] and [25].

The notable aspect among all the research on mean-variance problems with transaction costs is that simplifying assumptions are devised to solve the problem. Therefore more research is required into methods that can effectively solve non-convex problems with functions of the form (3.6). The starting point could be with the use of heuristic methods working on the problems

3.1.2 Cardinality Constrained Optimization Models

One of the modifications that can be made to the Markowitz model, is to add constraints on the number of assets to be held in the portfolio, called *cardinality constraints*.

Definition 3.1. The cardinality of $\mathbf{x}$ is defined as

$$Card(\mathbf{x}) = |\{i : x_i \neq 0\}|. \quad (3.13)$$

Clearly inclusion of the constraint (3.13) into the mean-variance problem leads to a non-convex problem. Suppose that exactly K assets are required in an optimal solution. The most natural way to incorporate (3.13) into the mean-variance is to convert it into a mixed integer quadratic programming (MIQP) problem as follows. Define the binary variable

$$z_i = \begin{cases} 1 & \text{if } x_i \neq 0 \\ 0 & \text{otherwise,} \end{cases} \quad (3.14)$$

and incorporate it into the constraint set of the mean-variance problem. This leads to the addition of the constraints

$$\begin{aligned} \sum_{i=1}^n z_i &= K \\ z_i l_i &\leq x_i \leq z_i u_i, \quad i = 1, \dots, n \\ z_i &\in \{0, 1\}, \quad i = 1, \dots, n. \end{aligned} \quad (3.15)$$

The cardinality constrained mean-variance (CCMV) problem is NP-complete [22]. When the cardinality constraint is an upper bound, that is $Card(\mathbf{x}) \leq K$, then it is a relaxation and the corresponding problem is generally easier to solve compared with that requiring $Card(\mathbf{x}) = K$. For a small number of assets n , the problem can be solved by state-of-the-art non-linear mixed integer solvers, like CPLEX. However, for a large number of assets, different methods have been suggested in literature to solve it. These methods can in general be grouped into three types: **exact algorithms**, **relaxation algorithms** and **heuristic algorithms**.

Exact Algorithms

These algorithms make a search within the feasible set and aim at finding an optimal solution. Most notable among such algorithms are the branch-and-bound algorithms and the

branch-and-cut algorithms.

Branch-and-Bound Methods for Cardinality Constrained Problems

In a branch-and-bound algorithm, the problem is subdivided into subproblems (usually two, but could be more). Then a relaxation of each subproblem is solved, sometimes to optimality and sometimes not, in order to determine or estimate the optimal objective function value of the relaxed subproblem. A subproblem can be eliminated from consideration if it is infeasible, **or** the solution to the subproblem has a higher objective function value than a known integer solution (assuming it is a minimization problem), **or** the solution to the subproblem satisfies the relaxed restrictions. If none of these cases holds, then a branching of the subproblem into new subproblems is done. This process forms a list of active problems. The process is repeated until no more branching can be done. Then the optimal solution is the best feasible solution that was encountered while solving the subproblems. Generally, branch-and-bound algorithms differ in terms of the branching criterion, how to choose an active subproblem and how to obtain a lower bound on the optimal cost of a subproblem.

Borchers and Mitchell [14] apply a branch-and-bound algorithm to solve the CCMV problem in which Lagrangian duality is used to obtain lower bounds on the optimal cost of the subproblems. Depth-first-strategy is used to choose an active subproblem until an integer solution has been found. It then switches to the best bound strategy. Branching is done at the variable with value closest to 0.5. They also employ early branching which helps to reduce on the computation time. In [15], they show that the method compares well against outer approximation algorithms, as it could solve problems which the outer approximation algorithm could not.

In [13], the variable to branch on is based on two new rules: *idiosyncratic risk branching procedure*, in which the variable chosen for branching is the one with the highest priority or most fractional, and the priorities are set before, and *portfolio risk branching procedure*, which chooses the variable to branch on whose integer feasibility restoration has the largest impact on the variance.

A CCMV problem with benchmarks and transaction costs is handled in [8] using a branch-and-bound algorithm which uses Lemke's pivoting method to solve the quadratic programming relaxation of the subproblem at each branch-and-bound node. Their algorithm uses the depth-first-search strategy to choose an active subproblem and branches on the variable with maximum absolute value first. Initialization of the upper bound of the objective value is done using re-optimization heuristics.

Branch-and-cut algorithms

These can be seen as a variant of the branch-and-bound method. At each node, one or several valid inequalities (cuts) is added to the problem. Branch-and-cut algorithms differ in terms of the method used to generate the valid inequalities, the branching method used, and how to choose active subproblems.

Bienstock [11] uses a branch-and-cut algorithm to solve a CCMV problem, which uses disjunctive procedure to generate the valid inequalities. As a rule of thumb, a cut is only used if the scaled violation is at least 10^{-3} . The method uses best node strategy (the one farthest from its bounds) to choose the next node to branch on. The method is also

coupled with heuristics to obtain an initial upper bound on the objective value.

A more problem specific type of cuts, called perspective cuts, are used by [47] to solve the MIQP problem.

Relaxation Algorithms

The relaxation algorithms aim at giving lower bounds or upper bounds or both for the CCMV problem. The relaxation can either be done in the objective or in the constraints set.

Shaw et al. [127] use Lagrangian duality to obtain lower bounds on the optimal cost of the subproblems, while a branching variable is chosen among those that are not fixed in the subproblem, and heuristics based on local search are used to initialize the upper bound of the problem. The method is tested on problems with up to 500 assets and outperforms CPLEX, because CPLEX uses quadratic programming relaxation, which is weaker than the Lagrangian relaxation.

A local relaxation method is employed in [109], which involves partitioning the constraint set into smaller subsets and solving subproblems on those sets. The solutions obtained from such subproblems provide center points, so that the problem is resolved on a neighbourhood of that center point. The process is repeated to get better solutions.

Heuristic Algorithms

These are adapted to a particular problem type and utilize experience of the structure of the solution. They are very useful in solving large-scale problems but may only give solutions that are close to an optimal solution. Some of the different heuristics used in literature to solve MIQP problems include the following.

Local search method The main idea behind the local search method is to pick a feasible point, say x^1 , which is obtained randomly or using some special technique, and evaluate the objective function at that point. Then a search is made in a neighbourhood of x^1 for another point, say x^2 , with a lower objective function value. If such a point is found, then x^2 replaces x^1 and the process is repeated starting at x^2 . The process is repeated until no point in the neighbourhood with lower objective function value can be found. The last point to be picked then is a local optimum of the problem. Local search algorithms will always differ in the way the neighbourhood is defined.

Ehrgott et al. [42] solve an MIQP problem using a two-phase local search algorithm based on two neighbourhood structures. The algorithm is called two-phase local search because a local search is performed on both neighborhoods alternately.

Metaheuristics These are extensions of the local search methods. They modify the search so that it can move to better solutions more easily. Below are some of the most common metaheuristic approaches that are used to solve MIQPs resulting from cardinality constrained portfolio optimization problems.

- **Simulated Annealing (SA):** The underlying idea of SA originated from an algorithm to simulate a certain thermodynamics process. The idea is to start at an initial point,

say x^1 , and obtain the value of the objective function, say $f(x^1)$. Then a search is made in the neighbourhood of x^1 for a point x^2 with a lower cost, $f(x^2)$, and if it is found, then x^2 replaces x^1 and the search continues; otherwise x^2 is accepted with a probability that decreases with the difference $f(x^2) - f(x^1)$ and the number of iterations. This gives it an advantage over the local search method, because it cannot get stuck in a local optimum.

- **Tabu Search (TS):** The TS is similar to the SA except that the TS has a memory about the history of visited points in the search. At a given point, a set of solutions that have been visited in the recent past, before a certain convergence criterion (for example a fixed number of iterations, CPU time, etc) is satisfied, is stored in a tabu list. Even if a point with lower cost function than x^1 is not found in the neighborhood of x^1 , a point with lowest objective function value will be picked to replace x^1 . In most cases, a TS is enriched with rules that enhance diversity and intensification in the search process.
- **Evolutionary Algorithm (EA):** The mechanisms of these algorithms stem from biological theory of evolution. The most common among EA algorithms, in solving CCMV problems, is the Genetic Algorithm (GA). The GA involves generating an initial population (points, in optimization). The individuals of the initial population are then evaluated by the "survival of the fittest" concept in biology (fitness function in optimization). The best parents are then selected from the initial population. These are recombined to produce children (new points) with better traits, who replace some or all the population. The process is repeated on the children up to when a satisfactory population (a set of solutions) has been found. For more details on GA, see [22] and [130].

It has also been found that a combination of two or more heuristics can lead to more efficient algorithms. These are called *hybrid algorithms*.

In Chang et al. [22], the CCMV problem is solved using three heuristic algorithms: GA, TS, and SA. They test their approach using problems with up to 225 assets. Their heuristic approach performs better than available state-of-the-art softwares for solving CCMV problems.

In addition to the two-phase local search algorithm, [42] solve the CCMV problem using three more metaheuristics: GA, SA and TS. The neighbourhoods used in these heuristics are those also used in the two-phase local search algorithm. They showed that the two-phase local search algorithm performs well on hard instances but the GA outcompetes the SA and TS.

Modified versions of the GA, TS and SA to solve the CCMV problem are used in [141]. Their heuristics make use of subset optimization. A subset of the portfolio with a return in a given prescribed range is chosen and solved to optimality at each stage. The modified method gave satisfactory results in terms of computational time, when tested on problems of up to 1318 assets.

A hybrid search algorithm which is a combination of the local search and quadratic programming techniques is used in [39]. It is shown that the hybrid search algorithm is superior to state-of-the-art solvers and also compares well with past developed algorithms for the same problem, e.g. [22] and [34].

Another hybrid algorithm which combines evolutionary techniques, in particular GA, with quadratic programming is used in [106]. The relaxation of the problem is first solved using a standard quadratic solver and the combinatorial part of the problem is handled using the GA.

In [34], the CCMV problem is solved using the SA algorithm. Their algorithm is tested on problems with up to 151 assets and results are obtained in a reasonable time.

A number of other heuristics in the literature have also been used to solve portfolio optimization problems under cardinality constraints. For example [46] solve the CCMV problem using neural network heuristics. They test their algorithm on problems with up to 225 assets. They show that the neural network heuristics compare well with the SA, GA and TS algorithms.

Other heuristics for solving the CCMV problem have been suggested, see [20], [12], [60], [125], [53] and [123].

Multi-Objective Portfolio Optimization Approach Under Cardinality Constraints

The bi-objective Markowitz model with cardinality and transaction level constraints is

$$\begin{aligned} \min_{\mathbf{x}} \quad & \{\mathbf{x}^T \Sigma \mathbf{x}, -\boldsymbol{\mu}^T \mathbf{x}\} \\ \text{s.t.} \quad & \text{Card}(\mathbf{x}) = K \\ & \mathbf{x} \in \mathcal{S}. \end{aligned} \tag{3.16}$$

Like the single-objective model with cardinality and transaction level constraints, model (3.16) is a MIQP. Most methods in literature to solve (3.16) are quite similar to those for solving the single-objective Markowitz model with cardinality and transaction level constraints.

Using a hybrid algorithm that combines the multi-objective evolutionary algorithms with quadratic programming local search methods, a method which [137] call multi-objective memetic algorithm, is used to solve (3.16).

Branke et al. [16] use a hybrid algorithm which employs the multi-objective evolutionary algorithm on subsets of the problem. The critical line algorithm in [100] is run on each of the subsets to get a solution to the subproblem, called an envelope. The final solution is then a combination of the partial solutions.

In [4], model (3.16) is solved using three metaheuristic approaches: greedy search, SA, and the ant colony approach. They showed that the ant colony approach was superior to the other two methods.

Other multi-objective heuristics can be found in [2] and [28].

3.1.3 Robust Optimization in the Mean-Variance Model

The mean-variance model can be very sensitive to changes in input parameters (see [30], [9] and [10]). Robust optimization can be used as a remedy to such a problem. Robust optimization refers to finding solutions to given optimization problems with uncertainty on the inputs, like parameters and distributions, that will achieve good objective values for all, or most, realizations of the uncertain inputs. The idea behind robust optimization is to assume an uncertainty set for an input parameter, or a distribution, and then find

an optimal solution that is valid for the uncertainty set. A detailed treatment of robust optimization is given in [1].

Assume uncertainty in the mean vector $\boldsymbol{\mu}$ and the covariance matrix $\boldsymbol{\Sigma}$, and suppose that the uncertainty sets for $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ are $\mathbb{U}_{\boldsymbol{\mu}}$ and $\mathbb{U}_{\boldsymbol{\Sigma}}$, respectively. If we consider the worst-case realization of $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$, then the robust counterparts of (3.3), (3.4) and (3.5) are (3.17), (3.18) and (3.19) respectively.

$$\begin{aligned}
 \max_{\mathbf{x}} \min_{\boldsymbol{\mu} \in \mathbb{U}_{\boldsymbol{\mu}}} \boldsymbol{\mu}^T \mathbf{x} \quad & \min_{\mathbf{x}} \max_{\boldsymbol{\Sigma} \in \mathbb{U}_{\boldsymbol{\Sigma}}} \mathbf{x}^T \boldsymbol{\Sigma} \mathbf{x} \quad & \max_{\mathbf{x}} \min_{\boldsymbol{\mu} \in \mathbb{U}_{\boldsymbol{\mu}}, \boldsymbol{\Sigma} \in \mathbb{U}_{\boldsymbol{\Sigma}}} \boldsymbol{\mu}^T \mathbf{x} - \lambda \mathbf{x}^T \boldsymbol{\Sigma} \mathbf{x} \\
 \text{s.t. } \max_{\boldsymbol{\Sigma} \in \mathbb{U}_{\boldsymbol{\Sigma}}} \mathbf{x}^T \boldsymbol{\Sigma} \mathbf{x} \leq \sigma_P^2 \quad & \text{s.t. } \min_{\boldsymbol{\mu} \in \mathbb{U}_{\boldsymbol{\mu}}} \boldsymbol{\mu}^T \mathbf{x} \geq \mu_P \quad & \text{s.t. } \mathbf{x} \in \mathcal{S}. \\
 \mathbf{x} \in \mathcal{S} \quad & \mathbf{x} \in \mathcal{S} \quad & (3.19)
 \end{aligned}
 \tag{3.17} \tag{3.18}$$

What then remains is to study problems (3.17), (3.18) and (3.19) under different uncertainty sets $\mathbb{U} = \{(\boldsymbol{\mu}, \boldsymbol{\Sigma}) : \boldsymbol{\mu} \in \mathbb{U}_{\boldsymbol{\mu}}, \boldsymbol{\Sigma} \in \mathbb{U}_{\boldsymbol{\Sigma}}\}$.

The interval uncertainty sets

$$\mathbb{U}_{\boldsymbol{\mu}} = \{\boldsymbol{\mu} : \boldsymbol{\mu}^L \leq \boldsymbol{\mu} \leq \boldsymbol{\mu}^U\}, \quad \mathbb{U}_{\boldsymbol{\Sigma}} = \{\boldsymbol{\Sigma} : \boldsymbol{\Sigma}^L \leq \boldsymbol{\Sigma} \leq \boldsymbol{\Sigma}^U, \boldsymbol{\Sigma} \succeq 0\}, \tag{3.20}$$

where $\boldsymbol{\mu}^L$, $\boldsymbol{\mu}^U$ and $\boldsymbol{\Sigma}^L$, $\boldsymbol{\Sigma}^U$ are the extreme values of the intervals, are considered in [139]. The uncertainty sets (3.20), are said to be of “box type”. For any given $\lambda > 0$, an optimal solution $\mathbf{x}^*(\lambda)$ for problem (3.19) is also an optimal solution for (3.18) for $\mu_P = \min_{\boldsymbol{\mu} \in \mathbb{U}_{\boldsymbol{\mu}}} \boldsymbol{\mu}^T \mathbf{x}^*(\lambda)$.

Let us denote the objective function in (3.19) as

$$\psi_{\lambda}(\mathbf{x}, \boldsymbol{\mu}, \boldsymbol{\Sigma}) = \boldsymbol{\mu}^T \mathbf{x} - \lambda \mathbf{x}^T \boldsymbol{\Sigma} \mathbf{x}. \tag{3.21}$$

Then the optimal solutions of the pair of primal and dual problems,

$$\max_{\mathbf{x} \in \mathcal{S}} \min_{(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \in \mathbb{U}} \psi_{\lambda}(\mathbf{x}, \boldsymbol{\mu}, \boldsymbol{\Sigma}) \quad \text{and} \quad \min_{(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \in \mathbb{U}} \max_{\mathbf{x} \in \mathcal{S}} \psi_{\lambda}(\mathbf{x}, \boldsymbol{\mu}, \boldsymbol{\Sigma}),$$

are equal at a saddle-point of the function $\psi_{\lambda}(\mathbf{x}, \boldsymbol{\mu}, \boldsymbol{\Sigma})$. This allows one to reformulate problem (3.19) as a problem of finding a saddle-point of the function $\psi_{\lambda}(\mathbf{x}, \boldsymbol{\mu}, \boldsymbol{\Sigma})$. Then problem (3.19) becomes

$$\begin{aligned}
 & \text{find } \bar{\mathbf{x}} \in \mathcal{S} \text{ and } (\bar{\boldsymbol{\mu}}, \bar{\boldsymbol{\Sigma}}) \in \mathbb{U} \text{ such that} \\
 & \psi_{\lambda}(\mathbf{x}, \bar{\boldsymbol{\mu}}, \bar{\boldsymbol{\Sigma}}) \leq \psi_{\lambda}(\bar{\mathbf{x}}, \bar{\boldsymbol{\mu}}, \bar{\boldsymbol{\Sigma}}) \leq \psi_{\lambda}(\bar{\mathbf{x}}, \boldsymbol{\mu}, \boldsymbol{\Sigma}), \quad \mathbf{x} \in \mathcal{S}, (\boldsymbol{\mu}, \boldsymbol{\Sigma}) \in \mathbb{U}.
 \end{aligned} \tag{3.22}$$

Suppose the vector of asset returns is given by

$$\mathbf{r} = \boldsymbol{\mu} + \mathbf{V}^T \mathbf{f} + \boldsymbol{\varepsilon} \tag{3.23}$$

where $\boldsymbol{\mu} \in \Re^n$ is the vector of mean asset returns, $\mathbf{f} \sim \mathcal{N}(0, \mathbf{F}) \in \Re^m$ is a vector factors that drive the market, $\mathbf{V} \in \Re^{m \times n}$ is the matrix of factor loadings of the n assets, $\boldsymbol{\varepsilon} \sim \mathcal{N}(0, \mathbf{D})$ is the vector of residual return. If we assume that $\boldsymbol{\varepsilon}$ is independent of $\mathbf{f}$, then $\mathbf{r} \sim \mathcal{N}(\boldsymbol{\mu}, \mathbf{V}^T \mathbf{F} \mathbf{V} + \mathbf{D})$. The following uncertainty sets are proposed in [52].

- For the covariance matrix $\mathbf{D}$,

$$S_d = \{\mathbf{D} : \mathbf{D} = \text{diag}(\mathbf{d}), d_i \in [d_i^L, d_i^U], i = 1, \dots, n\}. \quad (3.24)$$

- The matrix of factor loadings $\mathbf{V}$ belongs to the elliptical uncertainty set

$$S_v = \{\mathbf{V} : \mathbf{V} = \mathbf{V}_0 + \mathbf{W}, \|\mathbf{W}_i\|_{\Sigma} \leq \rho_i, i = 1, \dots, n\} \quad (3.25)$$

where $\mathbf{W}_i$ is the i^{th} column of $\mathbf{W}$, and $\|\mathbf{w}\|_{\Sigma} = \sqrt{\mathbf{w}^T \Sigma \mathbf{w}}$.

- The mean returns vector $\boldsymbol{\mu}$ belongs to the uncertainty set

$$S_m = \{\boldsymbol{\mu} : \boldsymbol{\mu} = \boldsymbol{\mu}_0 + \boldsymbol{\varepsilon}, |\varepsilon_i| \leq \gamma_i, i = 1, \dots, n\}. \quad (3.26)$$

So the returns on the portfolio $\mathbf{x}$ is

$$\mathbf{R} = \mathbf{r}^T \mathbf{x} + \mathbf{f}^T \mathbf{V} \mathbf{x} + \boldsymbol{\varepsilon}^T \mathbf{x} \sim \mathcal{N}(\mathbf{x}^T \boldsymbol{\mu}, \mathbf{x}^T (\mathbf{V}^T \mathbf{F} \mathbf{V} + \mathbf{D}) \mathbf{x}). \quad (3.27)$$

The robust analog of (3.3) is (3.28) and that of (3.4) is (3.29) below

$$\begin{aligned} \min_{\mathbf{x}} \quad & \max_{\mathbf{V} \in S_v, \mathbf{D} \in S_d} \text{Var}(\mathbf{R}) & \max_{\mathbf{x}} \quad & \min_{\boldsymbol{\mu} \in S_m} \mathbf{E}(\mathbf{R}) \\ \text{s.t.} \quad & \min_{\boldsymbol{\mu} \in S_m} \mathbf{E}(\mathbf{R}) \geq \mu_P & \text{s.t.} \quad & \max_{\mathbf{V} \in S_v, \mathbf{D} \in S_d} \text{Var}(\mathbf{R}) \leq \sigma_P^2 \\ & \mathbf{x} \in \mathcal{S} & & \mathbf{x} \in \mathcal{S}. \end{aligned} \quad (3.28) \quad (3.29)$$

Using equation (3.27), problem (3.28) becomes

$$\begin{aligned} \min_{\mathbf{x}} \quad & \max_{\mathbf{V} \in S_v} \|\mathbf{V} \mathbf{x}\|_{\mathbf{F}}^2 + \mathbf{x}^T \mathbf{D}^U \mathbf{x} \\ \text{s.t.} \quad & \min_{\boldsymbol{\mu} \in S_m} \boldsymbol{\mu}^T \mathbf{x} \geq \mu_P \\ & \mathbf{x} \in \mathcal{S}, \end{aligned} \quad (3.30)$$

where $\mathbf{D}^U = \text{diag}(\mathbf{d}^U)$.

Introducing auxiliary variables h and δ , problem (3.30) becomes

$$\begin{aligned} \min_{\mathbf{x}} \quad & h + \delta \\ \text{s.t.} \quad & \max_{\mathbf{V} \in S_v} \|\mathbf{V} \mathbf{x}\|_{\mathbf{F}}^2 \leq h \\ & \mathbf{x}^T \mathbf{D}^U \mathbf{x} \leq \delta \\ & \min_{\boldsymbol{\mu} \in S_m} \boldsymbol{\mu}^T \mathbf{x} \geq \mu_P \\ & \mathbf{x} \in \mathcal{S}. \end{aligned} \quad (3.31)$$

Goldfarb and Iyengar [52] show that (3.31) can be converted into a second order cone program (SOCP) and that it is hence solvable using for example interior point algorithms.

Suppose that we take a sample of the assets returns $\mathbf{r}^1, \dots, \mathbf{r}^q$ of size q and a sample of the factor returns $\mathbf{f}^1, \dots, \mathbf{f}^q$. Then the linear model (3.23) becomes

$$r_i^t = \bar{\mu}_i + \sum_{j=1}^m V_{ji} f_j^t + \varepsilon_i^t, \quad i = 1, \dots, n, \quad \text{for each } t = 1, \dots, q. \quad (3.32)$$

Let $\mathbf{B} = (\mathbf{f}^1, \dots, \mathbf{f}^q) \in \Re^{m \times q}$ be the matrix of factor returns and define $y_i = (r_i^1, \dots, r_i^q)^T$, $\mathbf{A} = (\mathbf{e}^T, \mathbf{B}^T) \in \Re^{q \times (m+1)}$, where $\mathbf{e}$ is a vector of ones, $p_i = (\mu_i, V_{1i}, \dots, V_{mi})^T$, $\boldsymbol{\varepsilon}_i = (\varepsilon_i^1, \dots, \varepsilon_i^q)^T$, where $\varepsilon_i^t \sim \mathcal{N}(0, \sigma_i^2)$, $i = 1, \dots, n$ and $t = 1, \dots, q$. Then equation (3.32) becomes

$$y_i = \mathbf{A}p_i + \boldsymbol{\varepsilon}_i, \quad i = 1, \dots, n. \quad (3.33)$$

The least-squares estimate $\bar{p}_i$, of p_i , is $\bar{p}_i = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T y_i$ (see [95]) and the unbiased estimate, s_i^2 of σ_i^2 , be given by

$$s_i^2 = \frac{\|y_i - \mathbf{A}\bar{p}_i\|^2}{q - m - 1}, \quad \text{for } i = 1, \dots, n.$$

Lu [95] considers a "joint" ellipsoidal uncertainty set for $(\boldsymbol{\mu}, \mathbf{V})$ with ω -confidence level given by

$$S_{\mu, \nu} = \left\{ (\tilde{\mu}, \tilde{V}) : \sum_{i=1}^n \frac{(\tilde{p}_i - \bar{p}_i)^T (\mathbf{A}^T \mathbf{A}) (\tilde{p}_i - \bar{p}_i)}{s_i^2} \leq (m+1) \tilde{c}(\omega) \right\} \quad (3.34)$$

for some $\tilde{c}(\omega)$, where $\tilde{p}_i = (\tilde{\mu}_i, \tilde{V}_{1i}, \dots, \tilde{V}_{mi})^T$, $i = 1, \dots, n$. So problem (3.19) under the "joint" uncertainty set (3.34) becomes

$$\begin{aligned} \max_{\mathbf{x}} \quad & \min_{(\boldsymbol{\mu}, \mathbf{V}) \in S_{\mu, \nu}} \mathbf{E}(R) - \lambda \text{Var}(R) \\ \text{s.t.} \quad & \mathbf{x} \in \mathcal{S}. \end{aligned} \quad (3.35)$$

Lu [95] reformulates problem (3.35) as a cone programming problem, which can be solved efficiently. According to [95], there are two drawbacks associated with using the separable uncertainty sets (as used by [52]), which are:

- The probability, of the uncertainty parameter falling within the uncertainty set (actual confidence level), is unknown and can even be much higher than the desired one.
- They are fully or partially box-typed. So the resulting robust portfolios can be too conservative.

Lu [94] demonstrates computationally that the robust portfolio determined by solving problem (3.35) using the "joint" uncertainty set outperforms that of a similar problem with uncertainty set (3.25) and (3.26).

Consider problem (3.5) in which there are $(n-1)$ risky assets and one risk-free asset, and where the investor receives information about $(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ from J different experts. That is to say, the investors gets $(\boldsymbol{\mu}_j, \boldsymbol{\Sigma}_j)$ for $j = 1, \dots, J$. The investor's problem is then to maximize the minimum expected utility implied by the various experts:

$$\begin{aligned} \max_{\mathbf{x}} \quad & \min_j (\boldsymbol{\mu}_j^T \mathbf{x} - \lambda \mathbf{x}^T \boldsymbol{\Sigma}_j \mathbf{x}) \\ \text{s.t.} \quad & \mathbf{x} \in \mathcal{S}. \end{aligned} \quad (3.36)$$

By letting $\lambda = \frac{1}{2}\gamma$, problem (3.36) becomes

$$\begin{aligned} \max_{\mathbf{x}} \quad & \min_j (\boldsymbol{\mu}_j^T \mathbf{x} - \frac{1}{2}\gamma \mathbf{x}^T \boldsymbol{\Sigma}_j \mathbf{x}) \\ \text{s.t.} \quad & \mathbf{x} \in \mathcal{S}. \end{aligned} \quad (3.37)$$

It is shown in [97] that the investor's optimal solution, which is the solution to problem (3.37), is $\mathbf{x}^* = \frac{1}{\gamma} S^{-1} m$, where $S = \sum_{j=1}^J \alpha_j \Sigma_j$, $m = \sum_{j=1}^J \alpha_j \mu_j$, with α_j being constants satisfying $0 \leq \alpha_j \leq 1$ and $\sum_{j=1}^J \alpha_j = 1$. Moreover, the values α_j are independent of the risk aversion γ . For general $(\boldsymbol{\mu}_j, \Sigma_j)$, the active Kuhn-Tucker constraints of problem (3.37) are determined numerically, which helps to determine the α_j 's. For analysis using the loss function and the disappointment function, both when Σ is known and not known, see [97] for details. See [37], [115] and [55] for other robust portfolio optimization models.

3.1.4 Multi-Period Mean-Variance Optimization

Portfolio optimization problems in the 1960s to late 1990s were mainly considered as expected utility maximization problems for investors and most of the work on multi-period portfolio optimization before 2000, was made in this context. The first mean-variance multi-period model, in the form of the Markowitz model, was handled in [87], who used dynamic programming to solve the resulting multi-stage problem. The problem of multi-period mean-variance optimization leads to a model which is non-separable. The non-separability arises due to the fact that expectation fulfills the tower property but variance does not, that is to say, for a random variable X and a filtration (a filtration $\{\mathcal{F}_t\}_{t \geq 0}$ is an information set available at time t , with $\{\mathcal{F}_s\}_s \subset \{\mathcal{F}_t\}_t$ for all $s < t$),

$$E(E(X/\mathcal{F}_s)) = E(X/\mathcal{F}_t), \quad \forall s > t \quad \text{but} \quad \text{Var}(\text{Var}(X/\mathcal{F}_s)) \neq \text{Var}(X/\mathcal{F}_t), \quad \forall s > t.$$

Let us consider a portfolio with $(n+1)$ risky assets at each period $t = 0, 1, \dots, T$. Let W_t be the wealth of the investor at the beginning of period t , r_t^i be a random rate of return of the i^{th} asset at time period t , so that $\mathbf{r}_t = (r_t^0, r_t^1, \dots, r_t^n)^T$ is a random returns vector at time period t , and let x_t^i be the amount invested in the i^{th} asset at the beginning of period t , so that $\mathbf{x}_t = (x_t^0, x_t^1, \dots, x_t^n)^T$ is an amount vector at period t . Assume further that the returns $\mathbf{r}_t$ are independent and $E(\mathbf{r}_t) = (E(r_t^0), E(r_t^1), \dots, E(r_t^n))^T$ and the covariances of $\mathbf{r}_t$ are known. Taking the 0^{th} asset as a reference asset, the amount invested in the 0^{th} asset at the beginning of time period t is given by

$$x_t^0 = W_t - \sum_{i=1}^n x_t^i,$$

and the wealth dynamics is given by

$$W_{t+1} = x_t^0 r_t^0 + \sum_{i=1}^n r_t^i x_t^i = \left(W_t - \sum_{i=1}^n x_t^i \right) r_t^0 + \sum_{i=1}^n r_t^i x_t^i.$$

Let $\mathbf{P}_t = ((r_t^1 - r_t^0), \dots, (r_t^n - r_t^0))^T$, so that

$$W_{t+1} = W_t r_t^0 + \mathbf{P}_t \mathbf{x}_t, \quad t = 0, 1, \dots, T-1. \quad (3.38)$$

In relation to Markowitz's model (3.1), [87] proposed three different formulations of the multi-period mean-variance model

$$\begin{array}{lll}
\min_{\mathbf{x}} & \text{Var}(W_T) & \max_{\mathbf{x}} \quad E(W_T) \\
\text{s.t.} & E(W_T) \geq \mu_P & \text{s.t.} \quad \text{Var}(W_T) \leq \sigma_P^2 \\
& W_{t+1} = W_t r_t^0 + \mathbf{P}_t \mathbf{x}_t & W_{t+1} = W_t r_t^0 + \mathbf{P}_t \mathbf{x}_t \\
& t = 0, 1, \dots, T-1 & t = 0, 1, \dots, T-1
\end{array}
\quad \begin{array}{l}
\max_{\mathbf{x}} \quad E(W_T) - \lambda \text{Var}(W_T) \\
\text{s.t.} \quad W_{t+1} = W_t r_t^0 + \mathbf{P}_t \mathbf{x}_t \\
t = 0, 1, \dots, T-1.
\end{array}
\quad (3.41)$$

(3.39)

(3.40)

A multi-period portfolio policy $\mathbf{x}^* = (\mathbf{x}_0^*, \mathbf{x}_1^*, \dots, \mathbf{x}_{T-1}^*)$ is efficient if $E(W_T)|_{\mathbf{x}^*} \geq E(W_T)|_{\mathbf{x}}$ and $\text{Var}(W_T)|_{\mathbf{x}^*} \leq \text{Var}(W_T)|_{\mathbf{x}}$ for all possible $\mathbf{x}$, with at least one inequality strict.

It is known [87] that for a given λ^* , if $\mathbf{x}^*$ solves (3.41), then it should also solve (3.40) with $\sigma_P^2 = \text{Var}(W_T)|_{\mathbf{x}^*}$ and also solve (3.39) with $\mu_P = E(W_T)|_{\mathbf{x}^*}$. As a guide to solve (3.41), the relation

$$\lambda = \frac{\partial E(W_T)}{\partial \text{Var}(W_T)}, \quad (3.42)$$

should hold at an optimal solution.

The idea that [87] adopts to solve (3.41) is to find a “somehow” related tractable auxiliary problem, which is separable, and find conditions under which solutions to the auxiliary problem also solve (3.41). Since the objective function of (3.41) can be rewritten as

$$u(W_T) = E(W_T) - \lambda \text{Var}(W_T) = -\lambda E(W_T^2) + [\lambda E(W_T)]^2 + E(W_T),$$

[87] proposes the following auxiliary problem

$$\begin{array}{ll}
\max & -\lambda E(W_T^2) + \theta E(W_T) \\
\text{s.t.} & W_{t+1} = W_t r_t^0 + \mathbf{P}_t \mathbf{x}_t, \quad t = 0, 1, \dots, T-1,
\end{array} \quad (3.43)$$

for some parameter θ . First, note that for any feasible portfolio policy $\mathbf{x}$,

$$\frac{\partial u(W_T)}{\partial E(W_T)} = 1 + 2\lambda E(W_T)|_{\mathbf{x}}.$$

The problem (3.43) is both convex and separable, and can be solved using dynamic programming.

Theorem 3.2 ([87])

If $\mathbf{x}^*$ is an optimal solution of (3.41), then $\mathbf{x}^*$ solves (3.43) for $\theta = 1 + 2\lambda E(W_T)|_{\mathbf{x}^*}$.

Theorem 3.3 ([87])

If for any optimal θ^* , $\mathbf{x}^*$ is an optimal solution of problem (3.43), then a necessary condition for $\mathbf{x}^*$ to be an optimal solution of problem (3.41) is that

$$\theta^* = 1 + 2\lambda E(W_T)|_{\mathbf{x}^*}.$$

The auxiliary problem (3.43) can be solved analytically and expressions for a closed form solution of (3.41) can be derived by making use of Theorem 3.2 and 3.3. Using the solution to (3.41), solutions to (3.40) and (3.39) can be determined.

In general, if a utility function u is a function of $E(W_T)$ and $\text{Var}(W_T)$, then the multi-period problem becomes

$$\begin{aligned} \max \quad & u[E(W_T), \text{Var}(W_T)] \\ \text{s.t.} \quad & W_{t+1} = W_t r_t^0 + \mathbf{P}_t \mathbf{x}_t, \quad t = 0, 1, \dots, T-1. \end{aligned} \quad (3.44)$$

The requirement that an investor should be risk averse means that the utility function u , should be concave. This requirement, together with independence of returns lead to the following.

$$\frac{\partial u[E(W_T), \text{Var}(W_T)]}{\partial E(W_T)} > 0, \quad \frac{\partial u[E(W_T), \text{Var}(W_T)]}{\partial \text{Var}(W_T)} < 0, \quad \text{and } E(\mathbf{r}_t \mathbf{r}_t^T) \succ 0.$$

Using these assumptions, [87] derived analytical solutions to (3.44). Leippold et al. [81] extended the work of [87] by including liabilities. Suppose q_t^i is the i^{th} liability return at the beginning of period t , v_t^i is the amount invested in liability i at time t and L_t is the liability at the beginning of time period t , then following similar arguments that led to (3.38), the dynamics of the liability is

$$\begin{aligned} L_{t+1} &= q_t^0 L_t + \mathbf{q}_t^T \mathbf{v}_t, \quad t = 0, 1, \dots, T-1 \\ \text{where } \mathbf{q}_t &= [(q_t^1 - q_t^0), (q_t^2 - q_t^0), \dots, (q_t^n - q_t^0)]^T \quad \text{with } q_t^0 = L_t - \sum_{i=1}^n v_t^i. \end{aligned} \quad (3.45)$$

The concern of the asset liability manager is the surplus $S_T = W_T - L_T$ at the terminal point. Therefore problem (3.39) is modified to (3.46), (3.40) modifies to (3.47) and (3.41) modifies to (3.48) below.

$$\begin{array}{lll} \min_{\mathbf{x}_t} \quad \text{Var}(S_T) & \max_{\mathbf{x}_t} \quad E(S_T) & \max_{\mathbf{x}_t} \quad E(S_T) - \lambda \text{Var}(S_T) \\ \text{s.t.} \quad E(S_T) \geq \mu_P & \text{s.t.} \quad \text{Var}(S_T) \leq \sigma_R^2 & \text{s.t.} \quad (3.38), (3.45) \\ (3.38), (3.45) & (3.38), (3.45) & (3.48) \\ (3.46) & (3.47) & \end{array}$$

Using a tractable separable auxiliary problem, [81] obtain conditions under which solutions to the auxiliary problem solve (3.48), and obtain a closed form solution.

For the case of exogenous liabilities, [24] study problem (3.48) but relax the positive definiteness requirement of $E(\mathbf{r}_t \mathbf{r}_t^T)$, which implies that some of those matrices can be singular. They use orthogonal transformations and also end up with closed form solutions to problem (3.48).

If the return constraint in (3.46) is assumed to be an equality, then [144] approach the problem by solving the dual through the use of dynamic programming and also obtains closed form solutions to problem (3.46).

The work of [24] is improved by [85], by including a bankruptcy control. A bankruptcy is said to occur if the surplus S_t at any time period t falls below a predefined "disaster level" b_t . Thus the probability of bankruptcy at time $t = 1, 2, \dots, T$ is

$$\mathbb{P}(S_t \leq b_t, S_j > b_j, \quad j = 1, \dots, t-1). \quad (3.49)$$

If we use the fact that $\mathbb{P}(S_t \leq b_t, S_j > b_j, j = 1, \dots, t-1) \leq \mathbb{P}(S_t \leq b_t)$, and the Tchebycheff inequality on (3.49), we get

$$\mathbb{P}(S_t \leq b_t) \leq \frac{\text{Var}(S_t)}{[\mathbb{E}(S_t) - b_t]^2}. \quad (3.50)$$

By imposing that $\mathbb{P}(S_t \leq b_t) \leq \frac{\text{Var}(S_t)}{[\mathbb{E}(S_t) - b_t]^2} \leq \alpha_t$, for an $\alpha_t \in (0, 1)$, problem (3.48) modifies to

$$\begin{aligned} \max_{\mathbf{x}} \quad & \mathbb{E}(S_T) - \lambda \text{Var}(S_T) \\ \text{s.t.} \quad & \text{Var}(S_t) \leq \alpha_t [\mathbb{E}(S_t) - b_t]^2 \\ & (3.38), (3.45). \end{aligned} \quad (3.51)$$

By using an auxiliary tractable problem for the dual problem of (3.51), closed form solutions for (3.51) are obtained [85].

Another modification to the multi-period optimization problem in the mean-variance sense is to consider a case where the market can be in different "regimes" at different times [19]. Let Y_t be the state of the market at time period t and $Y = \{Y_t, t = 0, 1, \dots\}$ is a homogeneous Markov chain with state space $E = \{1, 2, \dots, n\}$ and transition matrix $Q = (Q_{i,j})$. So the wealth dynamics (cf (3.38)) is given by

$$W_{t+1} = W_t r_t^0(Y_t) + \mathbf{P}_t^T(Y_t) \mathbf{x}_t(Y_t), \quad t = 0, 1, \dots, T-1. \quad (3.52)$$

In the stochastic market, problem (3.39) modifies to (3.53), (3.40) modifies to (3.54) and (3.41) modifies to (3.55).

$$\begin{aligned} \min_{\mathbf{x}_t} \quad & \text{Var}_i(W_T) & \max_{\mathbf{x}_t} \quad & \mathbb{E}_i(W_T) & \max_{\mathbf{x}_t} \quad & \mathbb{E}_i(W_T) - \lambda \text{Var}_i(W_T) \\ \text{s.t.} \quad & \mathbb{E}_i(W_T) \geq \mu_P & \text{s.t.} \quad & \text{Var}_i(W_T) \leq \sigma_P^2 & \text{s.t.} \quad & (3.52) \\ (3.52) & & (3.52) & & & \\ \text{with initial market state } i & & \text{with initial market state } i & & \text{with initial market state } i. & \\ (3.53) & & (3.54) & & (3.55) \end{aligned}$$

Again problems (3.55), (3.54) and (3.53) are non-separable and [19] also propose a tractable auxiliary problem

$$\begin{aligned} \max_{\mathbf{x}} \quad & -\lambda \mathbb{E}_i(W_T^2) + \theta \mathbb{E}_i(W_T) \\ \text{s.t.} \quad & W_{t+1} = W_t r_t^0(Y_t) + \mathbf{P}_t^T(Y_t) \mathbf{x}_t(Y_t), \quad t = 0, 1, \dots, T-1 \\ & \text{with initial market state } i, \end{aligned} \quad (3.56)$$

which is separable and thus obtain closed form solutions to (3.55). For a general utility maximization problem, the objective in (3.56) becomes $u[\mathbb{E}_i(W_T), \text{Var}_i(W_T)]$. Closed form solutions are similarly obtained for the quadratic utility function and the coefficient of variation (the ratio of standard deviation to the mean).

If a constraint on bankruptcy $\mathbb{P}_i(W_t \leq b_t)$ is considered, then the use of Tchebyshev's inequality and an upper bound α_t on the probability ensures that

$$\mathbb{P}_i(W_t \leq b) \leq \frac{\text{Var}_i(W_t)}{[\mathbb{E}_i(W_t) - b_t]^2} \leq \alpha_t. \quad (3.57)$$

The mean-variance portfolio optimization problem with a constraint on bankruptcy in a stochastic market is thus [147],

$$\begin{aligned}
 & \max_{\mathbf{x}_t} E_i(W_T) - \lambda \text{Var}_i(W_T) \\
 & \text{s.t. } \text{Var}_i(W_t) \leq \alpha_t [E_i(W_t) - b_t]^2 \\
 & W_{t+1} = W_t r_t^0(Y_t) + \mathbf{P}_t(Y_t) \mathbf{x}_t(Y_t), \quad t = 0, 1, \dots, T-1 \\
 & \text{given that the initial market state is } i.
 \end{aligned} \tag{3.58}$$

Wei and Ye [147] obtain closed form solutions of the dual of (3.58) using the ideas of [87], i.e. a tractable auxiliary problem.

For more modifications of the multi-period mean-variance optimization model, see [33], [96], [32], [90], [31] and [145]. The underlying principle for solving all these problems is the same, that is, the use a tractable auxiliary problem.

3.2 Mean Absolute Deviation Model

Due to the computational difficulty associated with the Markowitz model, [68] and [78] introduced an alternative risk measure, which would allow large scale problems to be solved easily. The resulting model from [78] is a linear programming (LP) problem and thus requires less computational time and memory compared to (3.1).

Definition 3.2 ([78]). The mean absolute deviation (MAD) of the portfolio returns $\mathbf{R}$ is

$$\text{MAD}(\mathbf{R}) = E(|\mathbf{R} - E(\mathbf{R})|). \tag{3.59}$$

Notice that MAD is an L_1 risk function. The MAD is in general a non-convex, non-differentiable function. Also MAD is not a coherent risk measure.

Theorem 3.4 ([78])

If the portfolio returns are multivariate normally distributed then

$$\text{MAD}(\mathbf{R}) = \sqrt{\frac{2\text{Var}(\mathbf{R})}{\pi}}.$$

Theorem 3.4 states that minimizing variance, which is an L_2 risk function, is equivalent to minimizing MAD, if the returns are multivariate normally distributed. The proposed MAD model, according to [68] and [78], is

$$\min_{\mathbf{x}} E(|\mathbf{r}^T \mathbf{x} - \boldsymbol{\mu}^T \mathbf{x}|) \tag{3.60a}$$

$$\text{s.t. } \boldsymbol{\mu}^T \mathbf{x} \geq \mu_P \tag{3.60b}$$

$$\mathbf{x} \in \mathcal{S}. \tag{3.60c}$$

The distributions of the random variables $\mathbf{r}$ are not known a priori, but they can be simulated using available historical data. Let $\mathbf{r}_t = (r_{1t}, \dots, r_{nt})$ be the realization of $\mathbf{r} =$

$(r_1, \dots, r_n)$ during the period $t = 1, \dots, T$ and assume that $p_t = \mathbb{P}\{(r_1, \dots, r_n) = (r_{1t}, \dots, r_{nt})\}$ is known in advance, and that $E(\mathbf{r}_t) = \bar{\boldsymbol{\mu}}$. Then (3.59) becomes

$$\text{MAD}(\mathbf{x}) = \sum_{t=1}^T p_t |\mathbf{r}_t^T \mathbf{x} - \bar{\boldsymbol{\mu}}^T \mathbf{x}|, \quad (3.61)$$

which replaces the objective function (3.60a). Let y_t be the smallest number which satisfies $|\mathbf{r}_t^T \mathbf{x} - \bar{\boldsymbol{\mu}}^T \mathbf{x}| \leq y_t$ and $-|\mathbf{r}_t^T \mathbf{x} - \bar{\boldsymbol{\mu}}^T \mathbf{x}| \leq y_t$. Then model (3.60) can be written as an LP problem

$$\begin{aligned} \min \quad & \mathbf{p}^T \mathbf{y} \\ \text{s.t.} \quad & (\mathbf{r}_t^T - \bar{\boldsymbol{\mu}}^T) \mathbf{x} \leq y_t, \quad t = 1, \dots, T \\ & -(\mathbf{r}_t^T - \bar{\boldsymbol{\mu}}^T) \mathbf{x} \leq y_t, \quad t = 1, \dots, T \end{aligned} \quad (3.62)$$

(3.60b), (3.60c),

where $\mathbf{p} = (p_1, \dots, p_T)$ and $\mathbf{y} = (y_1, \dots, y_T)$. A reformulation of (3.62) which reduces the number of variables is proposed in [45], by introducing non-negative variables v_t and ω_t , which satisfy

$$y_t + (\mathbf{r}_t^T - \bar{\boldsymbol{\mu}}^T) \mathbf{x} = 2v_t, \quad y_t - (\mathbf{r}_t^T - \bar{\boldsymbol{\mu}}^T) \mathbf{x} = 2\omega_t, \quad v_t \geq 0, \quad \omega_t \geq 0. \quad (3.63)$$

Using (3.63) to eliminate y_t from (3.62) leads to an LP problem (3.64) with T less constraints, compared to the model (3.62).

$$\begin{aligned} \min \quad & \sum_{t=1}^T (v_t + \omega_t) \\ \text{s.t.} \quad & v_t - \omega_t - p_t (\mathbf{r}_t^T - \bar{\boldsymbol{\mu}}^T) \mathbf{x} = 0, \quad t = 1, \dots, T \\ & (3.60b), (3.60c) \\ & v_t \geq 0, \quad \omega_t \geq 0, \quad t = 1, \dots, T. \end{aligned} \quad (3.64)$$

Another reformulation of (3.62) with the same number of auxiliary constraints as in [45], but with fewer number of additional continuous variables, is provided in [21] and it is superior to both (3.62) and (3.64) in terms of computational time. The MAD model has some interesting properties as seen in [71] and the survey [70]. Modifications of MAD are given in [104].

3.2.1 MAD under Transaction Costs

If a transaction cost $\phi_i(x_i)$ is associated with the i^{th} asset, then the expected rate of return under transaction costs is $\sum_{i=1}^n (\mu_i x_i - \phi_i(x_i))$. The MAD efficient frontier under transaction costs is determined by solving

$$\begin{aligned} \max \quad & \sum_{i=1}^n (\mu_i x_i - \phi_i(x_i)) \\ \text{s.t.} \quad & \text{MAD}(\mathbf{x}) \leq \Omega \\ & \mathbf{x} \in \mathcal{S}, \end{aligned} \quad (3.65)$$

where Ω is an upper limit on MAD.

In general, transaction costs $\phi_i(x_i)$ take on the form (3.6). However, like for the mean-variance problem, the models based on MAD incorporate simplified versions of (3.6). Below are some of the assumptions on $\phi_i(x_i)$ used in MAD models.

- When transaction costs $\phi_j(x_j)$ are assumed to be concave, then (3.65) is a linearly constrained convex minimization problem, which is in [72] solved using a branch-and-bound algorithm. They test the model on data of up to 200 assets. They show that the model can be extended to piecewise concave transaction cost functions.
- When $\phi_j(x_j)$ is a d.c function, that is, a difference of two convex functions, then (3.65) becomes a d.c optimization problem, which can also solved using a branch-and-bound algorithm [74].
- When $\phi_j(x_j)$ are either piecewise linear concave or piecewise constant with several jumps, then (3.65) is a non-convex mixed integer optimization problem, which is in [76] solved using a specialized branch-and-bound algorithm that out-competes CPLEX version 7.1.

For other papers on MAD under transaction costs see, [73], [75] and [69]. Kim et al. [67] consider a MAD model with both transaction costs and minimum transaction lots, while [77] considers a more general model with cardinality constraints and propose an algorithm to solve the resulting mixed integer linear program.

3.2.2 A Robust MAD Model

Consider uncertainty in the expected returns $\boldsymbol{\mu}$. The uncertainty set

$$\mathbb{U}_{\boldsymbol{\mu}} = \{\bar{\boldsymbol{\mu}} : r_{jt}^L \leq r_{jt} \leq r_{jt}^U, r_j^L \leq \bar{\mu}_j \leq r_j^U, j = 1, \dots, n, t = 1, \dots, T\}, \quad (3.66)$$

is proposed in [92]. Note that the uncertainty set (3.66) is determined by the sample returns. The solution technique used by [92] is to determine the upper and lower bounds of the objective in (3.62) on the interval set (3.66). The lower bound of the objective value of (3.62) under the uncertainty set (3.66) is

$$\begin{aligned} V^L = \min_{\bar{\boldsymbol{\mu}} \in \mathbb{U}_{\boldsymbol{\mu}}} \min_{\mathbf{x}} \quad & \mathbf{p}^T \mathbf{y} \\ \text{s.t.} \quad & (\mathbf{r}_t^T - \bar{\boldsymbol{\mu}}^T) \mathbf{x} \leq y_t, \quad t = 1, \dots, T \\ & -(\mathbf{r}_t^T - \bar{\boldsymbol{\mu}}^T) \mathbf{x} \leq y_t, \quad t = 1, \dots, T \end{aligned} \quad (3.67)$$

(3.60b), (3.60c).

The upper bound is

$$\begin{aligned} V^U = \max_{\bar{\boldsymbol{\mu}} \in \mathbb{U}_{\boldsymbol{\mu}}} \min_{\mathbf{x}} \quad & \mathbf{p}^T \mathbf{y} \\ \text{s.t.} \quad & (\mathbf{r}_t^T - \bar{\boldsymbol{\mu}}^T) \mathbf{x} \leq y_t, \quad t = 1, \dots, T \\ & -(\mathbf{r}_t^T - \bar{\boldsymbol{\mu}}^T) \mathbf{x} \leq y_t, \quad t = 1, \dots, T \end{aligned} \quad (3.68)$$

(3.60b), (3.60c).

The case with uncertainty in the constraint set is considered in [105].

3.3 Value-at-Risk Model

The value-at-risk (VaR) is the maximum expected loss over a specified period of time at a given confidence level. For example, if company A has a monthly VaR of SEK 2M at a 99% confidence level, then it means that there is a 1% probability that company A will incur losses of more than SEK 2M during any given one month period, if market conditions do not change. Therefore, VaR provides an estimate of the risk exposure of a portfolio.

Mathematically, if L is a random loss variable, then at a confidence level $\alpha \in (0, 1)$, VaR should satisfy

$$\mathbb{P}[L \leq \text{VaR}] = \alpha. \quad (3.69)$$

Hence, VaR at a confidence level α can be re-defined as an α -quantile of the distribution $F_L(\cdot)$ of L as

$$\text{VaR}_\alpha(L) = \inf\{l \in \mathfrak{R} : F_L(l) \geq \alpha\}. \quad (3.70)$$

VaR can be expressed either as a percentage returns (see [54], [121]) or as a monetary value (see [18]). VaR can also be regarded as a function of the random losses, see for example [121], [51], [122], or as a function of the random returns, see for example [143], [142], [48]. So L can either be random losses or future returns. VaR is a non-convex function and is not a coherent risk measure (see [57]). As a measure of risk, VaR became very famous and important in the 1990s and it was adopted by many financial institutions. In fact, [61] states that the Basle Committee on Banking supervision announced in 1995 that capital adequacy requirements for commercial banks were to be based on VaR.

The methods to compute VaR are normally based on the assumptions made on the distribution of L . These can be grouped into three main types.

1. Parametric method. The assumption in this method is that L follows a parametric distribution and so the VaR is the α -quantile of the distribution of L . For example, if the returns $\mathbf{r}$ of n assets in portfolio $\mathbf{x}$ are assumed to be normally distributed with expectation $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$, then the VaR at a confidence level α is

$$\text{VaR}_\alpha(\mathbf{R}) = Z_\alpha \sqrt{\mathbf{x}^T \boldsymbol{\Sigma} \mathbf{x}} - \boldsymbol{\mu}^T \mathbf{x}, \quad (3.71)$$

where Z_α is the α -quantile of the standard normal distribution i.e. $Z_\alpha = \Phi^{-1}(\alpha)$. Other distributions for the returns can be assumed or approximated by for example a log-normal distribution or a t -distribution.

2. Non-parametric method. Under this, no assumption is made about any specific distribution of the random losses. Historical data is used to simulate future random losses, which are used to approximate the VaR.
3. Monte-Carlo simulations. This is sometimes called the semi-parametric method. A known distribution is used to generate scenarios for the random losses, which are in turn used to calculate the VaR.

For a more detailed study on how to compute VaR, see [66], [40], [61] and [62].

If r^* is the maximum allowed VaR and we let $L = \mathbf{r}^T \mathbf{x}$, then the mean-VaR problem of portfolio optimization can take on the forms.

$$\max_{\mathbf{x}} \quad \boldsymbol{\mu}^T \mathbf{x} \quad (3.72a) \qquad \min_{\mathbf{x}} \quad \text{VaR}_{\alpha}(L) \quad (3.73a)$$

$$\text{s.t.} \quad \text{VaR}_{\alpha}(L) \leq r^* \quad (3.72b) \qquad \text{s.t.} \quad \boldsymbol{\mu}^T \mathbf{x} \geq \mu_P \quad (3.73b)$$

$$\mathbf{x} \in \mathcal{S} \quad (3.72c) \qquad \mathbf{x} \in \mathcal{S}. \quad (3.73c)$$

A portfolio $\mathbf{x}$ is said to belong to the mean-VaR efficient frontier at a confidence level α if and only if no other portfolio with a higher expected rate of return and a lower VaR exists [54]. If (3.73b) holds with an equality and $\mathcal{S} = \{\mathbf{x} \in \mathfrak{R}^n \mid \mathbf{e}^T \mathbf{x} = 1\}$, then the resulting optimal portfolio in (3.73) is said to belong to the mean-VaR boundary. It should be noted that problems (3.72) and (3.73) may not necessarily be equivalent due to the non-convexity of VaR.

If we let $A = \mathbf{e}^T \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu}$, $B = \boldsymbol{\mu}^T \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu}$, $C = \mathbf{e}^T \boldsymbol{\Sigma}^{-1} \mathbf{e}$, $D = BC - A^2$, then [54] show that a minimum VaR portfolio at a confidence level α exists if $\alpha > \Phi(\frac{D}{C})$ and it will then always be a minimum variance portfolio.

If the holding period Δt is considered, then (3.71) modifies to

$$\text{VaR}_{\alpha}(\mathbf{R}) = Z_{\alpha} \sqrt{\mathbf{x}^T \boldsymbol{\Omega} \mathbf{x}} \sqrt{\Delta t} - \boldsymbol{\mu}^T \mathbf{x} \Delta t.$$

The mean-VaR boundary problem

$$\begin{aligned} \min_{\mathbf{x}} \quad & \text{VaR}_{\alpha}(\mathbf{R}) = Z_{\alpha} \sqrt{\mathbf{x}^T \boldsymbol{\Omega} \mathbf{x}} \sqrt{\Delta t} - \boldsymbol{\mu}^T \mathbf{x} \Delta t \\ \text{s.t.} \quad & \mathbf{x}^T \boldsymbol{\mu} = \mu_P \\ & \mathbf{x}^T \mathbf{e} = 1, \end{aligned} \quad (3.74)$$

is solved using the Lagrange multiplier method to obtain a closed form solution [128].

Using historical data from more than 20 years, [48] showed that historical VaR is a very irregular function and it is highly non-convex, and that historical VaR is non-differentiable in every local minima. These properties render the optimization problem (3.72) under historical VaR, a very hard problem to solve. [48] proposed a numerical method that involves approximating the historical VaR with a smoothed VaR function which is differentiable and convex, leading to a convex problem which is easily solvable by state-of-art-software. The resulting solution, which is an approximation, could then be improved by locally minimizing the true VaR function starting at the obtained approximate solution. For general returns distributions, problem (3.72) is non-convex. A mixed integer LP (MILP) formulation of problems (3.72) and (3.73) can be achieved as follows ([7] and [91]). Let r_L be a lower bound on returns in the market and consider T scenarios, for which at each scenario t , a binary variable y_t is 1 if $\sum_{j=1}^n x_j r_{jt} \geq r_L$ and 0 otherwise. Then the MILP reformulation of (3.72) is

$$\max_{\mathbf{x}} \quad \sum_{t=1}^T p_t \sum_{j=1}^n x_j r_{jt} \quad (3.75a)$$

$$\text{s.t.} \quad \sum_{j=1}^n x_j r_{jt} \geq r_L + (r^* - r_L) y_t, \quad t = 1, \dots, T \quad (3.75b)$$

$$\sum_{t=1}^T p_t (1 - y_t) \leq \alpha \quad (3.75c)$$

$$y_t \in \{0, 1\}, \quad t = 1, \dots, T \quad (3.75d)$$

$$\mathbf{x} \in \mathcal{S}. \quad (3.75e)$$

It is shown in [7] that (3.75) is NP-hard. Solutions to (3.75) can be obtained using state-of-the-art softwares. Lin [91] suggests an MILP reformulation for (3.73) as

$$\begin{aligned} \max_{\mathbf{x}} \quad & r^* \\ \text{s.t.} \quad & (3.75b), (3.75c), (3.75d), \text{ and } (3.75e). \end{aligned} \quad (3.76)$$

When the returns are assumed to follow a discrete distribution, [126] solve (3.72) using a branch-and-bound algorithm to obtain a globally optimal solution. Furthermore, if the scenarios have equal probabilities, i.e. $p_t = \frac{1}{T}$, then the VaR can be written as a difference of two convex functions, called conditional value-at-risk, CVaR (to be covered in next section), as

$$\text{VaR}_\alpha(L) = k\text{CVaR}_{\frac{k}{T}}(L) - (k-1)\text{CVaR}_{\frac{k-1}{T}}(L) \quad \text{where } k = \lfloor \alpha T \rfloor. \quad (3.77)$$

Substituting (3.77) into (3.72b) leads to a d.c optimization problem of (3.72), which [143] solve using a branch-and-bound algorithm, and using a convex algorithm in [142]. A local search method to solve problem (3.73) is proposed in [51]. Problem (3.73) can also be reformulated as an LP problem with linear complementarity constraints, for which the upper and lower bounds of the objective function can be obtained easily. Pang and Leyffer [129] then proceed to apply a branch-and-cut algorithm to solve the problem to global optimality. Due to the shortcomings of the VaR, especially its non-convexity and non-coherence [5], the new risk measure called conditional value-at-risk was considered.

3.4 Conditional Value-at-Risk Model

Conditional value-at-risk, CVaR, also known as expected shortfall (see [51]) or average value-at-risk (see [143]) or tail value-at-risk (see [5], [23]), is defined as the expected loss exceeding VaR. Mathematically, if L is the random loss variable, then CVaR at a confidence level α is defined as

$$\text{CVaR}_\alpha(L) = \mathbb{E}[L \mid L > \text{VaR}_\alpha(L)].$$

For example, if the loss variable is the negative of the expected returns, that is, $L = -\mathbf{x}^T \boldsymbol{\mu}$, and $p(\cdot)$ is the density of the returns, then [121] define CVaR at a confidence level α as

$$\begin{aligned} \text{CVaR}_\alpha(\mathbf{x}) &= \min_{\beta \in \mathbb{R}} \beta + \frac{1}{1-\alpha} \int_{\mathbf{r} \in \mathbb{R}^n} [-\boldsymbol{\mu}^T \mathbf{x} - \beta]^+ p(\mathbf{r}) d\mathbf{r} \\ \text{where } [t]^+ &= \max(t, 0). \end{aligned} \quad (3.78)$$

CVaR is a convex function and a coherent risk measure [122]. A theoretical comparison between CVaR and VaR is given in [114]. The mean-CVaR optimization problem is

$$\min_{\mathbf{x}} \quad \text{CVaR}_\alpha(\mathbf{x}) = \min_{\mathbf{x}, \beta} \beta + \frac{1}{1-\alpha} \int_{\mathbf{r} \in \mathbb{R}^n} [-\boldsymbol{\mu}^T \mathbf{x} - \beta]^+ p(\mathbf{r}) d\mathbf{r} \quad (3.79a)$$

$$\text{s.t.} \quad \boldsymbol{\mu}^T \mathbf{x} \geq \mu_P \quad (3.79b)$$

$$\mathbf{x} \in \mathcal{S}. \quad (3.79c)$$

The integral in the objective function in (3.79) can be approximated by selecting a sample of the returns vector, for example from the historical returns. Let $\mathbf{r}^1, \mathbf{r}^2, \dots, \mathbf{r}^q$ be our representative sample of vector $\mathbf{r}$. Then the objective function in (3.79) is approximated by

$$\beta + \frac{1}{q(1-\alpha)} \sum_{k=1}^q [-\mathbf{r}^k{}^T \mathbf{x} - \beta]^+. \quad (3.80)$$

We can then define auxiliary variables u_k , $k = 1, \dots, q$ such that $u_k \geq -\mathbf{r}^k{}^T \mathbf{x} - \beta$ and $u_k \geq 0$, so that the portfolio optimization problem (3.79) is approximated by

$$\min_{\mathbf{x}} \text{CVaR}_{\alpha}(\mathbf{x}) = \min_{\mathbf{x}, \beta, \mathbf{u}} \beta + \frac{1}{q(1-\alpha)} \sum_{k=1}^q u_k \quad (3.81a)$$

$$\text{s.t.} \quad u_k \geq -\mathbf{r}^k{}^T \mathbf{x} - \beta \quad k = 1, \dots, q \quad (3.81b)$$

$$(3.79b), (3.79c),$$

$$u_k \geq 0, \quad k = 1, \dots, q. \quad (3.81c)$$

Problem (3.81) is an LP problem and can be solved using any LP solver. (3.81) gives the same optimal solution as the mean-variance model, if the returns are normally distributed. The value of β which solves problem (3.81) is in fact the VaR. The discretization which was done to come up with (3.80) does not take into account the nature of the probability distribution of the returns $\mathbf{r}$, since it assumes that $\mathbb{P}(r_i) = 1/q$, $i = 1, \dots, n$.

Suppose that we introduce general probabilities p_k , of the scenarios $\mathbf{r}^k$ in the objective function of (3.79), then (3.81) becomes the more general problem

$$\min_{\mathbf{x}} \text{CVaR}_{\alpha}(\mathbf{x}) = \min_{\mathbf{x}, \beta, \mathbf{u}} \beta + \frac{1}{(1-\alpha)} \sum_{k=1}^q p_k u_k \quad (3.82)$$

$$\text{s.t.} \quad (3.81b), (3.81c), (3.79b), (3.79c).$$

An alternative formulation of the CVaR optimization problem is to maximize expected returns $\bar{\boldsymbol{\mu}}^T \mathbf{x}$ subject to CVaR constraints. If ω is the maximum allowed CVaR value, then the problem becomes

$$\max_{\mathbf{x}} \quad \bar{\boldsymbol{\mu}}^T \mathbf{x}$$

$$\text{s.t.} \quad \beta + \frac{1}{(1-\alpha)} \sum_{k=1}^q p_k u_k \leq \omega \quad (3.83)$$

$$(3.81b), (3.81c), (3.79b), (3.79c).$$

Krokhmal et al. [79] show that problems (3.82) and (3.83) will generate the same efficient frontier as long as $\bar{\boldsymbol{\mu}}^T \mathbf{x} \geq \mu_P$ and $\text{CVaR}_{\alpha}(\mathbf{x}) \leq \omega$ have interior points.

Transaction Costs and Other Constraints

Let us consider a single-period model so that $\mathbf{x}^0 = (x_1^0, \dots, x_n^0)^T$ represents our initial portfolio holdings, $\mathbf{x} = (x_1, \dots, x_n)^T$ represents the final optimal portfolio that we intend to find, $\mathbf{r}^0 = (r_1^0, \dots, r_n^0)^T$ represents the vector of initial returns on the portfolio and $\mathbf{r} = (r_1, \dots, r_n)^T$ represents the vector of final returns on the portfolio (which are unknown). Then the return over the period is $\mathbf{r}^{0T} \mathbf{x}^0 - \mathbf{r}^T \mathbf{x}$. Consider a linear proportional transaction cost f_i , such that when buying or selling the i^{th} asset, one pays f_i times the amount of transaction. A balance constraint that maintains the total returns on the portfolio including transaction costs is

$$\sum_{i=1}^n r_i^0 x_i^0 = \sum_{i=1}^n f_i r_i^0 |x_i^0 - x_i| + \sum_{i=1}^n r_i^0 x_i. \quad (3.84)$$

If we let $|x_i^0 - x_i|$ be the smallest number m_i which satisfies $x_i^0 - x_i \leq m_i$ and $-x_i^0 + x_i \leq m_i$, then (3.84) can be re-written as

$$\begin{aligned} \sum_{i=1}^n r_i^0 x_i^0 &= \sum_{i=1}^n f_i r_i^0 m_i + \sum_{i=1}^n r_i^0 x_i \\ m_i &\geq x_i^0 - x_i, \quad m_i \geq -x_i^0 + x_i. \end{aligned} \quad (3.85)$$

Other constraints in the optimization problem can be added, for example:

- Constraints on the returns,

$$r_i^0 x_i \leq \rho_i \sum_{k=1}^n r_k^0 x_k, \quad i = 1, \dots, n, \quad (3.86)$$

where ρ_i is percentage.

- Bounds on portfolio positions,

$$l_i \leq x_i \leq u_i, \quad i = 1, \dots, n. \quad (3.87)$$

The optimization problem with constraints and transaction costs becomes a modification of (3.83) when (3.85), (3.86) and (3.87) are added, leading to

$$\begin{aligned} \max_{\mathbf{x}, \beta, \mathbf{u}} \quad & \bar{\boldsymbol{\mu}}^T \mathbf{x} \\ \text{s.t.} \quad & \beta + \frac{1}{(1-\alpha)} \sum_{k=1}^q p_k u_k \leq \omega \\ & u_k \geq -(\mathbf{r}^{0T} \mathbf{x}^0 - \mathbf{r}^{kT} \mathbf{x}) - \beta, \quad u_k \geq 0, \quad k = 1, \dots, q \\ & \text{constraints (3.85), (3.86), (3.87), (3.81c).} \end{aligned} \quad (3.88)$$

Problem (3.88) was studied in [79]. It is an LP problem and can be solved efficiently.

Addition of both fixed transaction costs and cardinality constraints into problem (3.88), leads to an NP-hard problem [3].

3.4.1 Robust Optimization Using CVaR

Consider uncertainty in different model inputs.

Uncertainty in returns. Assume that the expected returns belong to an uncertainty set $\mathbb{U}_\mu$. If we consider worst case returns, then the robust counterpart of (3.81) is

$$\begin{aligned} \min_{\mathbf{x}} \quad & \text{CVaR}_\alpha(\mathbf{x}) = \min_{\mathbf{x}, \beta, \mathbf{u}} \quad \beta + \frac{1}{q(1-\alpha)} \sum_{k=1}^q p_k \\ \text{s.t.} \quad & \min_{\bar{\boldsymbol{\mu}} \in \mathbb{U}_\mu} \bar{\boldsymbol{\mu}}^T \mathbf{x} \geq \mu_P \end{aligned} \quad (3.89)$$

(3.81b), (3.81c), (3.79b), (3.79c).

Solving problem (3.89) requires one to specify the geometry of the uncertainty set $\mathbb{U}_\mu$.

Quaranta and Zaffaroni [119] used what they termed as the ‘‘Soyster’s approach’’, based on [131]. The expected returns μ_i are known within confidence regions, with a confidence level around their estimates $\hat{\mu}_i$, leading to

$$\mathbb{U}_\mu = \{\bar{\boldsymbol{\mu}} = (\mu_1, \dots, \mu_n) : \hat{\mu}_i - s_i \leq \mu_i \leq \hat{\mu}_i + s_i, \quad i = 1, \dots, n\}.$$

Then

$$\min_{\bar{\boldsymbol{\mu}} \in \mathbb{U}_\mu} \bar{\boldsymbol{\mu}}^T \mathbf{x} = \sum_{i=1}^n \hat{\mu}_i x_i - \sum_{i=1}^n s_i |x_i|.$$

If we assume that $|x_i| \leq b_i$, then problem (3.89) becomes

$$\begin{aligned} \min_{\mathbf{x}} \quad & \text{CVaR}_\alpha(\mathbf{x}) = \min_{\mathbf{x}, \beta, \mathbf{u}} \quad \beta + \frac{1}{q(1-\alpha)} \sum_{k=1}^q p_k \\ \text{s.t.} \quad & \sum_{i=1}^n \hat{\mu}_i x_i - \sum_{i=1}^n s_i b_i \geq \mu_P \end{aligned} \quad (3.90)$$

(3.81b), (3.81c), (3.79b), (3.79c).

Problem (3.90) is an LP problem.

Uncertainty in returns distribution. When the probability distribution of returns $p(\mathbf{r})$ is known to belong to an uncertainty set $\mathbb{U}_{p(\mathbf{r})}$ of distributions, [148] define the worst-case CVaR (WCVaR) with respect to $\mathbb{U}_{p(\mathbf{r})}$ as

$$\text{WCVaR}_\alpha(\mathbf{x}) = \sup_{p(\mathbf{r}) \in \mathbb{U}_{p(\mathbf{r})}} \text{CVaR}_\alpha(\mathbf{x}). \quad (3.91)$$

Since CVaR is a coherent risk measure, then even WCVaR is also a coherent risk measure [148]. Minimization of WCVaR will depend on the uncertainty set $\mathbb{U}_{p(\mathbf{r})}$. The work by [148] considers three uncertainty structures: mixture distribution uncertainty, box uncertainty and ellipsoidal uncertainty, which lead to an LP problem, an LP problem, and an SOCP problem, respectively. For details see [148] and [43].

If the distribution p belongs to the set of asset price distributions at maturity that replicate current prices of options on the assets, then [59] propose the uncertainty set for a distribution $p \in \mathfrak{R}_+$ as

$$\mathbb{U}_p = \{p : E(1) = 1, E(\mathbf{S}_T) = \mathbf{p}^0, E[(\mathbf{S}_T - \mathbf{K}^i)^+] = \mathbf{p}^i, i = 1, \dots, n\},$$

where $\mathbf{S}_T$ is a vector of uncertain prices at maturity time T , $\mathbf{p}^0$ is the price vector of a European forward option on assets $1, \dots, n$ maturing at time T , and $\mathbf{p}^i$ is the price vector of a European call option on assets $1, \dots, n$ with strike price $\mathbf{K}^i$, maturing at time T . If all the assets are assumed to be arbitrage free, then [59] showed that the resulting WCVaR problem can be converted into an LP problem.

If the distribution p belongs to the uncertainty set

$$\mathbb{U}_p = \{E(\mathbf{r}) = \boldsymbol{\mu}, \text{Cov}(\mathbf{r}) = \Sigma \succ 0\},$$

with just a simple constraint $\mathbf{e}^T \mathbf{x} = 1$, the analytical solution to the WCVaR problem is given in [86].

For more uncertainty sets and how to construct risk measures, the reader is referred to [110].

3.5 Mean Absolute Semi Deviation Model

In relation to the MAD suggested by [78] and [68], another closely related risk measure called mean absolute deviation (MASD), was proposed in [133].

Definition 3.3 ([133]). The MASD of a portfolio $\mathbf{x}$ is defined as

$$\text{MASD}(\mathbf{x}) = E(\max\{\boldsymbol{\mu}^T \mathbf{x} - \mathbf{r}^T \mathbf{x}, 0\}).$$

Speranza [133] showed that optimization using MAD is equivalent to using MASD. In fact the later is a half of the former. Using similar scenarios like for the MAD, we define MASD as

$$\text{MASD}(\mathbf{x}) = \sum_{t=1}^T p_t \min\{0, \sum_{j=1}^n (r_{jt} - \bar{\mu}_j) x_j\},$$

which can be re-written as

$$\begin{aligned} \min \quad & \mathbf{p}^T \mathbf{y} \\ \text{s.t.} \quad & y_t + \sum_{j=1}^n (r_{jt} - \bar{\mu}_j) x_j \geq 0, \quad t = 1, \dots, T \\ & y_t \geq 0 \quad t = 1, \dots, T. \end{aligned} \tag{3.92}$$

The MASD portfolio optimization problem thus becomes the solution of (3.92) subject to (3.60b) and (3.60c).

3.5.1 MASD with Real Features

In this section, we explain the modifications made to the MASD portfolio optimization model, when real features are added

Transaction costs

Let f_i be the proportional transaction cost associated with the i^{th} asset and F_i be the corresponding fixed transaction cost, and assume that both fixed and proportional transaction costs are simultaneously incurred on an asset. Let z_i be defined as in (3.14). Then the return constraint $\sum_{j=1}^n \mu_j x_j \geq \mu_P$ modifies to

$$\sum_{j=1}^n (\mu_j - d_j) z_j x_j - \sum_{j=1}^n p_j z_j \geq \mu_P \quad (3.93)$$

$$z_j \in \{0, 1\}, \quad j = 1, \dots, n.$$

Cardinality constraints

If K is the maximum number of assets required in a portfolio, then the cardinality constraint is

$$\sum_{j=1}^n z_j \leq K. \quad (3.94)$$

Minimum transaction lots

Let m_j be the fraction of the total funds required to purchase a minimum lot of the j^{th} asset and g_j be the number of minimum lots for the j^{th} asset. Then the constraint is

$$\sum_{j=1}^n m_j g_j = C, \quad (3.95)$$

where $C \in (0, 1]$ is the total portfolio expenditure of available funds.

Kellerer et al. [65] show that a MASD model with (3.93) is NP-hard and proposes a heuristic to solve the problem. Heuristics to solve problems with real features are proposed in [134] and [98]. An exact algorithm which involves partitioning the feasible set into smaller partitions and solving the problem on each of the partitions is proposed by [99]. A closely related problem for mutual funds is handled by [29].

Concluding Remarks

We have reviewed the major mean-risk models in the literature since 1952 up to date. Even with the emergence of new mean-risk models, more research has been done on the mean-variance model compared with the other mean-risk models.

There is still a great need to devise new and more effective solution techniques to handle mean-risk models with real features like transaction costs, cardinality constraints and others. When real features are incorporated into the models, the resulting models are usually non-convex and hard to solve.

Bibliography

- [1] Aharon Ben-Tal, L. E. G. and Nemirovski, A. (2009). *Robust Optimization*. Princeton University Press.
- [2] Anagnostopoulos, K. P. and Mamanis, G. (2011). The mean-variance cardinality constrained portfolio optimization problem: An experimental evaluation of five multi-objective evolutionary algorithms. *Expert Systems with Applications*, 38(11) 14208–14217.
- [3] Angelelli, E., Mansini, R., and Speranza, M. G. (2008). A comparison of mad and cvar models with real features. *Journal of Banking & Finance*, 32(7):1188–1197.
- [4] Armañanzas, R. and Lozano, J. A. (2005). A multiobjective approach to the portfolio optimization problem. In *Congress on Evolutionary Computation*, 2:1388–1395.
- [5] Artzner, P., Delbaen, F., Eber, J. M., and Heath, D. (1999). Coherent Measures of Risk. *Mathematical Finance*, 9(3): 203–228.
- [6] Balbás, A. (2007). Mathematical methods in modern risk measurement: a survey. Open Access publications from Universidad Carlos III de Madrid 2, Universidad Carlos III de Madrid.
- [7] Benati, S. and Rizzi, R. (2007). A mixed integer linear programming formulation of the optimal mean/value-at-risk portfolio problem. *European Journal of Operational Research*, 176(1):423–434.
- [8] Bertsimas, D. and Shioda, R. (2009). Algorithm for cardinality-constrained quadratic optimization. *Computational Optimization and Applications*, 43(1):1–22.
- [9] Best, M. J. and Grauer, R. R. (1991). Sensitivity analysis for mean-variance portfolio problems. *Management Science*, 37(8):980–989.

- [10] Best, M. J. and Grauer, R. R. (1992). Positively weighted minimum-variance portfolios and the structure of asset expected returns. *Journal of Financial and Quantitative Analysis*, 27(04):513–537.
- [11] Bienstock, D. (1995). Computational study of a family of mixed-integer quadratic programming problems. *Mathematical programming*, 74(2):121–140.
- [12] Blog, B., van der Hoek, G., Kan, A. H. G. R., and Timmer, G. T. (1983). The optimal selection of small portfolios. *Management Science*, 29(7):792–798.
- [13] Bonami, P. and Lejeune, M. A. (2009). An exact solution approach for portfolio optimization problems under stochastic and integer constraints. *Operations Research*, 57(3):650–670.
- [14] Borchers, B. and Mitchell, J. E. (1994). An improved branch and bound algorithm for mixed integer nonlinear programs. *Computers and Operations Research*, 21(4):359–367.
- [15] Borchers, B. and Mitchell, J. E. (1997). A computational comparison of branch and bound and outer approximation algorithms for 0-1 mixed integer nonlinear programs. *Computers and Operations Research*, 24(8):699–701.
- [16] Branke, J., Scheckenbach, B., Stein, M., Deb, K., and Schmeck, H. (2009). Portfolio optimization with an envelope-based multi-objective evolutionary algorithm. *European Journal of Operational Research*, 199(3):684–693.
- [17] Cai, X., Teo, K.-L., Yang, X., and Zhou, X. Y. (2000). Portfolio optimization under a minimax rule. *Management Science*, 46(7):957–972.
- [18] Campbell, R., Huisman, R., and Koedijk, K. (2001). Optimal portfolio selection in a value-at-risk framework. *Journal of Banking & Finance*, 25:1789–1804.
- [19] Celikyurt, U. and Ozekici, S. (2007). Multiperiod portfolio optimization models in stochastic markets using the mean-variance approach. *European Journal of Operational Research*, 179(1):186–202.
- [20] Cesarone, F., Scozzari, A., and Tardella, F. (2012). A new method for mean-variance portfolio optimization with cardinality constraints. *Annals of Operations Research*, pages 1–22.
- [21] Chang, C.-T. (2005). A modified goal programming approach for the mean-absolute deviation portfolio optimization model. *Applied Mathematics and Computation*, 171(1):567–572.
- [22] Chang, T.-J., Meade, N., Beasley, J. E., and Sharaiha, Y. M. (2000). Heuristics for cardinality constrained portfolio optimisation. *Computers and Operations Research*, 27(13):1271–1302.
- [23] Charpentier, A. and Oulidi, A. (2009). Estimating allocations for value-at-risk portfolio optimization. *Mathematical Methods of Operation Research*, 69(3):395–410.

- [24] Chen, P., Yang, H., and Yin, G. (2008). Markowitz's mean-variance asset-liability management with regime switching: A continuous-time model. *Insurance: Mathematics and Economics*, 43(3):456–465.
- [25] Chen, W. and Zhang, W.-G. (2010). The admissible portfolio selection problem with transaction costs and an improved pso algorithm. *Physica A: Statistical Mechanics and its Applications*, 389(10):2070 – 2076.
- [26] Cherny, A. S. (2009). Capital allocation and contribution with discrete-time coherent risk. *Mathematical Finance*, 19(1):13–40.
- [27] Cherny, A. S. (2010). Risk-reward optimization with discrete-time coherent risk. *Mathematical Finance*, 20(4):571–595.
- [28] Chiam, S., Tan, K., and Al Mamum, A. (2008). Evolutionary multi-objective portfolio optimization in practical context. *International Journal of Automation and Computing*, 5:67–80. 10.1007/s11633-008-0067-2.
- [29] Chiodi, L., Mansini, R., and Speranza, M. G. (2003). Semi-absolute deviation rule for mutual funds portfolio selection. *Annals of Operations Research*, 124(1-4):245–265.
- [30] Chopra, V. K., Hensel, C. R., and Turner, A. L. (1993). Massaging mean-variance inputs: returns from alternative global investment strategies in the 1980s. *Management Science*, 39(7):845–855.
- [31] Costa, O. and Nabholz, R. (2007). Multiperiod mean-variance optimization with intertemporal restrictions. *Journal of Optimization Theory and Applications*, 134:257–274.
- [32] Costa, O. and Okimura, R. (2009). Discrete-time mean variance optimal control of linear systems with markovian jumps and multiplicative noise. *International Journal of Control*, 82(2):256–267.
- [33] Costa, O. L. V. and Araujo, M. V. (2008). A generalized multi-period mean-variance portfolio optimization with markov switching parameters. *Automatica*, 44(10):2487–2497.
- [34] Crama, Y. and Schyns, M. (2003). Simulated annealing for complex portfolio selection problems. *European Journal of Operational Research*, 150(3):546–571.
- [35] Dantzig, G. B. and Infanger, G. (1993). Multi-stage stochastic linear programs for portfolio optimization. *Annals of Operations Research*, 45(1):59–76.
- [36] Delbaen, F. (2000). *Coherent Risk Measures on General Probability Spaces*. Springer, Berlin.
- [37] DeMiguel, V. and Nogales, F. J. (2009). Portfolio selection with robust estimation. *Operations Research*, 57(3):560–577.
- [38] Denault, M. (2001). Coherent allocation of risk capital. *Journal of Risk*, 4:1–34.

- [39] Di Gaspero, L., di Tollo, G., Roli, A., and Schaerf, A. (2011). Local search for constrained financial portfolio selection problems with short sellings. In *Proceedings of the 5th International Conference on Learning and Intelligent Optimization, LION'05*, 6683: 450–453, Berlin, Heidelberg. Springer-Verlag.
- [40] Duffie, D. and Pan, J. (1997). An overview of value at risk. *The Journal of Derivatives*, 4(3):7–49.
- [41] Eeckhoudt, L., Gollier, C., and Schlesinger, H. (2005). *Economic and Financial Decisions under Risk*. Princeton University Press.
- [42] Ehrgott, M., Klamroth, K., and Schwehm, C. (2004). An mcdm approach to portfolio optimization. *European Journal of Operational Research*, 155(3):752–770.
- [43] Fabozzi, F. J., Huang, D., and Zhou, G. (2010). Robust portfolios: contributions from operations research and finance. *Annals of Operations Research*, 176(1):191–220.
- [44] Fama, E. F. (1970). Multiperiod consumption-investment decisions. *American Economic Review*, 60(1):163–74.
- [45] Feinstein, C. D. and Thapa, M. N. (1993). A reformulation of a mean-absolute deviation portfolio optimization model. *Management Science*, 39(12):1552–1553.
- [46] Fernandez, A. and Gomez, S. (2007). Portfolio selection using neural networks. *Computers and Operations Research*, 34(4):1177–1191.
- [47] Frangioni, A. and Gentile, C. (2006). Perspective cuts for a class of convex 0-1 mixed integer programs. *Mathematical Programming*, 106(2):225–236.
- [48] Gaivoronski, A. A. and Pflug, G. (2005). Value-at-risk in portfolio optimization: properties and computational approach. *Journal of Risk*, 7(2):1–31.
- [49] Gaspero, L. D., Tollo, G. D., Roli, A., and Schaerf, A. (2010). Hybrid metaheuristics for constrained portfolio selection problems. *Quantitative Finance*, 11(10):1–15.
- [50] Gennotte, G. and Jung, A. (1994). Investment strategies under transaction costs: The finite horizon case. *Management Science*, 40(3):385–404.
- [51] Gilli, M., Kllezi, E., and Hysi, H. (2006). A data-driven optimization heuristic for downside risk minimization. *The Journal of Risk*, 8(3):1–18.
- [52] Goldfarb, D. and Iyengar, G. (2003). Robust portfolio selection problems. *Mathematics of Operations Research*, 28(1):1–38.
- [53] Gomez, M. A., Flores, C. X., and Osorio, M. A. (2006). Hybrid search for cardinality constrained portfolio optimization. In *Proceedings of the 8th annual conference on Genetic and evolutionary computation, GECCO '06*, 1865–1866, New York, NY, USA. ACM.

- [54] Gordon J. A., Alexander, M. B. (2002). Economic implications of using a mean-var model for portfolio selection: A comparison with mean-variance analysis. *Journal of Economic Dynamics & Control*, 26(7-8):1159–1193.
- [55] Gulpinar, N., An, L. T. H., and Moeini, M. (2010). Robust investment strategies with discrete asset choice constraints using dc programming. *Optimization*, 59(1):45–62.
- [56] Hakansson, N. H. (1970). Optimal investment and consumption strategies under risk for a class of utility functions. *Econometrica*, 38(5):587–607.
- [57] Henrion, R. (2006). Some remarks on value-at-risk optimization. *International Journal of Management Science and Engineering Management*, 1(2):111–118.
- [58] Hirschberger, M., Qi, Y., and Steuer, R. E. (2010). Large-scale mv efficient frontier computation via a procedure of parametric quadratic programming. *European Journal of Operational Research*, 204:581–588.
- [59] Jabbour, C., na, J. F. P., Vera, J. C., and Zuluaga, L. F. (2008). An estimation-free, robust conditional value-at-risk portfolio allocation model. *The Journal of Risk*, 11(1):57–78.
- [60] Jobst, N. J., Horniman, M. D., Lucas, C. A., and Mitra, G. (2001). Computational aspects of alternative portfolio selection models in the presence of discrete asset choice constraints. *Quantitative Finance*, 1(5):489–501.
- [61] Jorion, P. (1996). Risk2: Measuring the risk in value at risk. *Financial Analysts Journal*, 52(6):47–56.
- [62] Jorion, P. (2007). *Value at Risk: The New Benchmark for Managing Financial Risk*. McGraw-Hill.
- [63] Jorion, P. (2009). *Financial Risk Manager Handbook*. Wiley finance series.
- [64] Kawadai, N. and Konno, H. (2001). Solving large scale mean-variance models with dense non-factorable covariance matrices. *Journal of the Operations Research Society of Japan*, 44(3):251–279.
- [65] Kellerer, H., Mansini, R., and Speranza, M. (2000). Selecting portfolios with fixed costs and minimum transaction lots. *Annals of Operations Research*, 99:287–304.
- [66] Khindanova, I. N. and Rachev, S. T. (2000). Value at risk: Recent advances. In *Handbook on Analytic-Computational Methods in Applied Mathematics*. CRC Press LLC.
- [67] Kim, J. S., Kim, Y. C., and Shin, K. Y. (2005). An algorithm for portfolio optimization problem. *Informatica*, 16(1):93–106.
- [68] Konno, H. (1990). Piecewise linear risk function and portfolio optimization. *Journal of the Operations Research Society of Japan*, 33(2):139–156.

- [69] Konno, H., Akishino, K., and Yamamoto, R. (2005). Optimization of a long-short portfolio under non-convex transaction cost. *Computational Optimization and Applications*, 32(1-2):115–132.
- [70] Konno, H. and Koshizuka, T. (2005). Mean-absolute deviation model. *IIIE Transactions*, 37(10):893–900.
- [71] Konno, H. and Shirakawa, H. (1994). Equilibrium relations in a capital asset market: A mean absolute deviation approach. *Asia-Pacific Financial Markets*, 1:21–35. 10.1007/BF02425207.
- [72] Konno, H. and Wijayanayake, A. (1999). Mean-absolute deviation portfolio optimization model under transaction costs. *Journal of the Operations Research Society of Japan*, 42(4):422–435.
- [73] Konno, H. and Wijayanayake, A. (2001). Portfolio optimization problem under concave transaction costs and minimal transaction unit constraints. *Mathematical Programming*, 89:233–250. 10.1007/PL00011397.
- [74] Konno, H. and Wijayanayake, A. (2002). Portfolio optimization under dc. transaction costs and minimal transaction unit constraints. *Journal of Global Optimization*, 22(1-4):137–154.
- [75] Konno, H. and Yamamoto, R. (2003). Minimal concave cost rebalance of a portfolio to the efficient frontier. *Mathematical Programming*, 97(3):571–585.
- [76] Konno, H. and Yamamoto, R. (2005a). Global optimization versus integer programming in portfolio optimization under nonconvex transaction costs. *Journal of Global Optimization*, 32(2):207–219.
- [77] Konno, H. and Yamamoto, R. (2005b). Integer programming approaches in mean-risk models. *Computational Management Science*, 2:339–351.
- [78] Konno, H. and Yamazaki, H. (1991). Mean-absolute deviation portfolio optimization model and its applications to tokyo stock market. *Management Science*, 37(5):519–531.
- [79] Krokmal, P., Palmquist, J., and Uryasev, S. (2002). Portfolio optimization with conditional value-at-risk objective and constraints. *Journal of Risk*, 4:11–27.
- [80] Kroll, Y., Levy, H., and Markowitz, H. M. (1984). Mean-variance versus direct utility maximization. *Journal of Finance*, 39(1):47–61.
- [81] Leippold, M., Trojani, F., and Vanini, P. (2004). A geometric approach to multiperiod mean variance optimization of assets and liabilities. *Journal of Economic Dynamics and Control*, 28(6):1079–1113.
- [82] Lemrabott, M., Gueye, S., Yassine, A., and Rakotondratsimba, Y. (2008). Portfolio selection under piecewise affine transaction costs: An integer quadratic formulation. In Le Thi, H., Bouvry, P., and Pham Dinh, T., editors, *Modelling, Computation and*

- Optimization in Information Systems and Management Sciences*, volume 14 of *Communications in Computer and Information Science*, pages 190–196. Springer Berlin Heidelberg.
- [83] Levy, H. and Markowitz, H. M. (1979). Approximating expected utility by a function of mean and variance. *The American Economic Review*, 69(3):308–317.
- [84] Li, C. and Li, Z. (2012a). Multi-period portfolio optimization for asset-liability management with bankrupt control. *Applied Mathematics and Computation*, 218(22):11196–11208.
- [85] Li, C. and Li, Z. (2012b). Multi-period portfolio optimization for asset-liability management with bankrupt control. *Applied Mathematics and Computation*, 218(22):11196–11208.
- [86] Li, C., Simai, H., and Shuzhong, Z. (2011). Tight bounds for some risk measures, with applications to robust portfolio selection. *Operations Research*, 59(4):847–865.
- [87] Li, D. and Ng, W.-L. (2000). Optimal dynamic portfolio selection: Multiperiod mean-variance formulation. *Mathematical Finance*, 10(3):387–406.
- [88] Li, D., Sun, X., and Wang, J. (2006). Optimal lot solution to cardinality constrained mean-variance formulation for portfolio selection. *Mathematical Finance*, 16(1):83–101.
- [89] Li, Z.-F., Wang, S.-Y., and Deng, X.-T. (2000). A linear programming algorithm for optimal portfolio selection with transaction costs. *International Journal of Systems Science*, 31(1):107–117.
- [90] Liang, J. (2009). Discrete analysis of portfolio selection with optimal stopping time. *Journal of Applied Mathematics and Decision Sciences*, 2009.
- [91] Lin, C.-C. (2009). Comments on A mixed integer linear programming formulation of the optimal mean/Value-at-Risk portfolio problem. *European Journal of Operational Research*, 194(1):339 – 341.
- [92] Liu, S.-T. (2011). The mean-absolute deviation portfolio selection problem with interval-valued returns. *Journal of Computational and Applied Mathematics*, 235(14):4149–4157.
- [93] Lobo, M. S., Fazel, M., and Boyd, S. (2007). Portfolio optimization with linear and fixed transaction costs. *Annals of Operation Research*, 152(1):341–365.
- [94] Lu, Z. (2011a). A computational study on robust portfolio selection based on a joint ellipsoidal uncertainty set. *Mathematical Programming*, 126(1):193–201.
- [95] Lu, Z. (2011b). Robust portfolio selection based on a joint ellipsoidal uncertainty set. *Optimization Methods and Software*, 26(1):89–104.
- [96] Luo, S. (2007). Multi-period asset allocation under hidden markovianly driven noises. *Stochastic Analysis and Applications*, 25(5):1057–1078.

- [97] Lutgens, F. and Schotman, P. C. (2010). Robust portfolio optimisation with multiple experts. *Review of Finance*, 14(2):343–383.
- [98] Mansini, R. and Speranza, M. G. (1999). Heuristic algorithms for the portfolio selection problem with minimum transaction lots. *European Journal of Operational Research*, 114(2):219–233.
- [99] Mansini, R. and Speranza, M. G. (2005). An exact approach for portfolio selection with transaction costs and rounds. *IIIE Transactions*, 37(10):919–929.
- [100] Markowitz, H. (1952). Portfolio Selection. *The Journal of Finance*, 7(1):77–91.
- [101] Markowitz, H. M. (1959). *Portfolio Selection: Efficient Diversification of Investments*. John Wiley & Sons.
- [102] McNeil, A. J., Frey, R., and Embrechts, P. (2005). *Quantitative Risk Management: Concepts, Techniques and Tools*. Princeton University Press.
- [103] Merton, R. C. (1972). An analytic derivation of the efficient portfolio frontier. *The Journal of Financial and Quantitative Analysis*, 7(4): 1851–1872.
- [104] Michalowski, W. and Ogryczak, W. (1998). Extending the mad portfolio optimization model to incorporate downside risk aversion. Working Papers ir98041, International Institute for Applied Systems Analysis.
- [105] Moon, Y. and Yao, T. (2011). A robust mean absolute deviation model for portfolio optimization. *Computers and Operations Research*, 38(9):1251–1258.
- [106] Moral-Escudero, R., Ruiz-Torrubiano, R., and Suarez, A. (2006). Selection of optimal investment portfolios with cardinality constraints. In *Evolutionary Computation, 2006. CEC 2006. IEEE Congress on*, pages 2382–2388.
- [107] Mossin, J. (1968). Optimal multi-period portfolio policies. *The Journal of Business*, 41(2):215–229.
- [108] Mulvey, J. M. and Vladimirov, H. (1992). Stochastic network programming for financial planning problems. *Management Science*, 38(11):1642–1664.
- [109] Murray, W. and Shek, H. (2012). A local relaxation method for the cardinality constrained portfolio optimization problem. *Computational Optimization and Applications*, 53(3):681–709.
- [110] Natarajan, K., Pachamanova, D., and Sim, M. (2009). Constructing risk measures from uncertainty sets. *Operations Research*, 57(5):1129–1141.
- [111] Nocedal, J. and Wright, S. J. (2006). *Numerical Optimization*. Springer, New York, 2nd edition.
- [112] Patel, N. R. and Subrahmanyam, M. G. (1982). A simple algorithm for optimal portfolio selection with fixed transaction costs. *Management Science*, 28(3):303–314.

- [113] Perold, A. F. (1984). Large-scale portfolio optimization. *Management Science*, 30(10):1143–1160.
- [114] Pflug, G. C. (2000). *Some Remarks on the Value-at-Risk and the Conditional value-at-risk.*, volume 1 of *Probabilistic Constrained Optimization*, section 1.2, pages 272–281. Kluwer, second edition.
- [115] Popescu, I. (2007). Robust mean-covariance solutions for stochastic optimization. *Operations Research*, 55(1):98–112.
- [116] Potapchik, M., Tunçel, L., and Wolkowicz, H. (2008). Large scale portfolio optimization with piecewise linear transaction costs. *Optimization Methods and Software*, 23(6):929 – 952.
- [117] Pratt, J. W. (1964). Risk aversion in the small and in the large. *Econometrica*, 32(1/2):122–136.
- [118] Qi, Y., Wang, Z., and Shen, P. (2010). Can covariance matrix refinements alleviate the contradiction of mean-variance efficiency and diversification of portfolio selection? In *Management and Service Science (MASS), 2010 International Conference on*, pages 1–4.
- [119] Quaranta, A. G. and Zaffaroni, A. (2008). Robust optimization of conditional value at risk and portfolio selection. *Journal of Banking & Finance*, 32(10):2046–2056.
- [120] Rockafellar, R., Uryasev, S., and Zabarankin, M. (2006). Generalized deviations in risk analysis. *Finance and Stochastics*, 10(1):51–74.
- [121] Rockafellar, R. T. and Uryasev, S. (2000). Optimization of conditional value-at-risk. *Journal of Risk*, 2:21–41.
- [122] Rockafellar, R. T. and Uryasev, S. (2002). Conditional value-at-risk for general loss distributions. *Journal of Banking & Finance*, 26(7):1443–1471.
- [123] Rubén, R.-T. and Suárez, A. (2010). Hybrid approaches and dimensionality reduction for portfolio selection with cardinality constraints. *Computational Intelligence Magazine, IEEE*, 5(2):92 –107.
- [124] Ryoo, H. S. (2007). A compact mean-variance-skewness model for large-scale portfolio optimization and its application to the nyse market. *The Journal of the Operational Research Society*, 58(4):505–515.
- [125] Schaerf, A. (2002). Local search techniques for constrained portfolio selection problems. *Computational Economics*, 20(3):177–90.
- [126] Cheon, M.S, Ahmed, S., and Al-khayyal, F. (2007). A branch-reduce-cut algorithm for the global optimization of probabilistically constrained linear programs. *Mathematical Programming*, 108(2-3):617–634
- [127] Shaw, D. X., Liu, S., and Kopman, L. (2008). Lagrangian relaxation procedure for cardinality-constrained portfolio optimization. *Optimization Methods and Software*, 23(3):411–420.

- [128] Sheng, Z., Benshana, S., and Zhongping, W. (2012). Analysis of mean-var model for financial risk control. *Systems Engineering Procedia*, 4:40–45.
- [129] Pang, J. and Leyffer, S. (2004). On the global minimization of the value-at-risk. *Optimization Methods and Software*, 19(5):611–631.
- [130] Soleimani, H., Golmakani, H. R., and Salimi, M. H. (2009). Markowitz-based portfolio selection with minimum transaction lots, cardinality constraints and regarding sector capitalization using genetic algorithm. *Expert Systems and Applications*, 36(3):5058–5063.
- [131] Soyster, A. L. (1973). Convex programming with set-inclusive constraints and applications to inexact linear programming. *Operations Research*, 21(5):pp. 1154–1157.
- [132] Speranza, M. G. (1993a). Linear programming models for portfolio optimization. *Finance*, 14:107–123.
- [133] Speranza, M. G. (1993b). Linear programming models for portfolio optimization. *Finance*, 14:107–123.
- [134] Speranza, M. G. (1996). A heuristic algorithm for a portfolio optimization model applied to the milan stock market. *Computers and Operations Research*, 23(5):433 – 441.
- [135] Stein, M., Branke, J., and Schmeck, H. (2008). Efficient implementation of an active set algorithm for large-scale portfolio selection. *Computers and Operations Research*, 35:3945 – 3961.
- [136] Steuer, R. E., Qi, Y., and Hirschberger, M. (2011). Comparative issues in large-scale mean-variance efficient frontier computation. *Decision Support Systems*, 51:250–255.
- [137] Streichert, F., Ulmer, H., and Zell, A. (2003). Evolutionary algorithms and the cardinality constrained portfolio optimization problem. In *Operations Research Proceedings 2003, Selected Papers of the International Conference on Operations Research OR 2003*, pages 253–260. Springer.
- [138] Tobin, J. (1958). Liquidity preference as behavior towards risk. *Review of Economic Studies*, 25(2):65–86.
- [139] Tütüncü, R. H. and Koenig, M. (2004). Robust asset allocation. *Annals of Operations Research*, 132(1-4):157–187.
- [140] von Neumann, J. and Morgenstern, O. (1994). *Theory of Games and Economic Behavior*. Princeton University Press.
- [141] Woodside-Oriakhi, M., Lucas, C., and Beasley, J. (2011). Heuristic algorithms for the cardinality constrained efficient frontier. *European Journal of Operational Research*, 213(3):538 – 550.

- [142] Wozabal, D. (2012). Value-at-risk optimization using the difference of convex algorithm. *OR Spectrum*, 34:861–883.
- [143] Wozabal, D., Hochreiter, R., and Pflug, G. C. (2010). A difference of convex formulation of value-at-risk constrained optimization. *Optimization*, 59(3):377–400.
- [144] Yao, H. (2011). A simple method for solving multiperiod mean-variance asset-liability management problem. *Procedia Engineering*, 23:387 – 391.
- [145] Yao, H., Zheng, H., Ma, Q., and Ma, Y. (2010). Multi-period mean-variance model with uncertain exit time. In *Information Management, Innovation Management and Industrial Engineering (ICIII), 2010 International Conference on*, volume 2, pages 43–46.
- [146] Xue, H.-G., Xu, C.-X., and Feng, Z.-X. (2006). Mean-variance portfolio optimal problem under concave transaction cost. *Applied Mathematics and Computation*, 174(1):1 – 12.
- [147] Wei, S. and Ye, Z. (2007). Multi-period optimization portfolio with bankruptcy control in stochastic market. *Applied Mathematics and Computation*, 186(1):414 – 425.
- [148] Zhu, S. and Fukushima, M. (2009). Worst-case conditional value-at-risk with application to robust portfolio management. *Operations Research*, 57(5):1155–1168.

Part II

Research Papers

Research Papers

The articles associated with this thesis have been removed for copyright reasons. For more details about these see:

<http://urn.kb.se/resolve?urn=urn:nbn:se:liu:diva-118362>