Solutions to VaR / CVaR Optimization Problems for Decision Making under Uncertainty

Graduate School of Systems and Information Engineering

University of Tsukuba

March 2010

Yuichi Takano

Abstract

It is unrealistic to formulate the problems arising under uncertain environments as deterministic optimization problems which assume that their modeling parameters were given and fixed. Uncertainty should be taken into consideration. Uncertainty can also invite a risk of incurring a unexpectedly large loss. Risk measure would be crucial to develop a useful optimization technique to avert the risk. Among most preferable risk measures in financial risk management are the Value-at-Risk (VaR) and the Conditional Value-at-Risk (CVaR). We consider three kinds of stochastic programming problems where the VaR or CVaR is minimized for attaining an optimal decision.

Firstly, we consider the CVaR minimization in the context of the well-known newsvendor problem, which is originally formulated as the maximization of the expected profit. Under mild assumptions on the probability distribution of demand for product, we provide an analytical solution or a linear programming formulation for the case where the analytical solution is not available. Numerical examples are also exhibited, clarifying the difference among the models analyzed in this thesis, and demonstrating the efficiency of the linear programming solutions.

Secondly, we study the VaR minimization with a finite number of scenarios. The nonconvex feasible set makes the problem intractable, so that we propose to approximate the feasible set by the difference of two convex sets (D.C. set), a new conservative approximation. We show that it is formulated as a D.C. optimization problem whose complexity is independent of the number of scenarios. We propose a simplicial branch-and-bound algorithm for a solution of the problem, and report some numerical experiments.

Thirdly, we study the constant rebalancing strategy for a multi-period portfolio optimization with transaction cost subject to CVaR constraints. In general this problem is difficult to solve due to the nonconvexity. Nonlinearity of the transaction cost and CVaR constraint makes things worse, and a locally optimal solution may not be reached via a state-of-the-art nonlinear programming solver. As a practical solution, we propose a local search algorithm where linear approximation problems and nonlinear equations are iteratively solved. Computational results are presented, showing that the proposed algorithm attains a good solution in practical time.

Contents

List of Figures

List of Tables

Notations

Throughout this thesis, we use the following notations.

notation	definition and explanation
$\mathbb{R}$	set of real numbers
$\mathbb{R}^n$	set of n dimensional real column vectors
$\mathbb{R}_+$	set of nonnegative real numbers
$\boldsymbol{a}^\top$, $\boldsymbol{A}^\top$	transposed vector of $\boldsymbol{a}$, transposed matrix of $\boldsymbol{A}$
$\lvert A \rvert$	cardinality of set A
$\min A$	minimum number of set A
$\max A$	maximum number of set A
$\mathbf{0}$	column vector of zeros of an appropriate dimension
$\mathbf{1}$	column vector of ones of an appropriate dimension
$\mathbb{E}[\tilde{X}]$	expected value of random variable $\tilde{X}$
$\mathrm{Prob}\, F$	probability of event F
$1\!\!1_A$	indicator function of set A, i.e., $1\!\!1_A(x) := \begin{cases} 1 & \text{if } x \in A \\ 0 & \text{if } x \notin A \end{cases}$
$[\ \ ]^+$	plus function, i.e., $[\,x\,]^+ := \max\{0, x\}$
exp	exponential function
log	logarithm function
$\mathrm{int}\, A$	interior of set A
∇f	gradient of function f

Chapter 1

Introduction

Mathematical formulation of real world problems invariably includes uncertain parameters. Only their stochastic properties are available from empirical data or are anticipated as a mathematical assumption, hence it would pose a risk of making irrelevant and even wrong decisions to assume inattentively that those parameters were given and fixed. This fact makes stochastic programming an absolutely vital approach to optimization in diverse areas. In this chapter, we briefly explain stochastic programming and discuss several modeling techniques and approaches to a solution.

1.1 Motivation for Stochastic Programming

The research field of optimization has remarkably developed for decades since Dantzig's monumental work [24]. We are now able to solve large scale optimization problems efficiently owing to the improvement of optimization algorithms and their implementation techniques as well as the recent progress in computer technologies. Observing phenomena in the real world, finding problems, formulating as optimization problems and then solving for an optimal solution is now a very helpful engineering process for decision makers of central and local governments, companies of various kinds. Optimization problems thus constructed usually include parameters that are uncertain by nature, such as demand of inventory, return of financial assets and the like. However we often neglect the unavoidable uncertainty inherent in those modeling parameters and assume that their values are fixed and known in advance. This is clearly not a practical method because the optimal solution to be obtained is unlikely to be an efficient solution. Thus the problems should be solved with uncertainty of parameters into consideration, i.e., as a stochastic programming problem. There have been proposed a good many stochastic models and methods (e.g., [41, 76]). Among them we will survey three stochastic programming problems in the subsequent section: *News Vendor Problem*, *Probabilistically Constrained Problem* and *Multi-Period Portfolio Optimization Problem*.

1.2 Background, Purpose and Results of the Thesis

In decision making in the presence of uncertainty, avoiding unacceptably large loss should be given the first priority. A natural question would be what risk measure is appropriate to evaluate the large loss. Among various risk measures proposed so far, the *Value-at-Risk* (VaR) and the *Conditional Value-at-Risk* (CVaR) should be best suited to this purpose. In this thesis we will apply these two risk measures to three different stochastic programming problems and develop a solution method for each. We explain the background, purpose and results for the three problems in the following subsections.

1.2.1 Newsvendor Solutions via Conditional Value-at-Risk Minimization

Suppose that a manager has to place an order every day for a product that is subject to uncertain demand and becoming almost worthless on the next day. This situation is well formulated as the classic newsvendor model. It offers a solution that maximizes the daily expected profit or minimizes the daily expected cost. However, looking at the daily expected profit alone is not satisfactory from a practical viewpoint because it does not pay attention to the risk of potential and possibly large loss. An alternative scheme is to attain a predetermined target profit in a somewhat high probability, but this scheme may still result in an unacceptably large loss. To reduce such a risk, some researchers have proposed to minimize the standard deviation of profit (e.g., [50]), which originates from the mean-variance model of Markowitz [53].

The profit below the target level is a risk to hedge, but that above the target level is not. Therefore, minimizing a downside risk measure that captures a risk of the profit going down to some target level is more appealing than minimizing the standard deviation. For example, Lau [50] and Lau and Lau [49] examine a model which maximizes the probability of exceeding a predetermined fixed target profit, whereas Parlar and Weng [67] consider the expected profit in place of the fixed target profit. These objectives are intuitively comprehensible, but their maximization leads to a nonconvex optimization problem, and accordingly, their objectives are very difficult to achieve for general distribution functions.

We propose to adopt the *conditional value-at-risk* (CVaR), which has been widely used recently in the area of inventory management (e.g., [1, 9, 21, 60]), as the downside risk measure. The CVaR is a risk measure with preferable properties, namely, it is coherent ([4]) and consistent with the second order stochastic dominance ([69, 65]). These properties are known to be equivalent to some axiomatization of rational investors' behavior under uncertainty. In particular, the consistency with the stochastic dominance implies that minimizing the CVaR never conflicts with maximizing the expectation of any risk-averse utility function ([65]). Some researchers, e.g., Eeckhoudt et al. [28], treat the risk-aversion directly through the newsvendor's utility function. However, utility function is too conceptual and difficult to identify in practice, so that the use of risk measures has advantage over that of utility functions.

In financial portfolio management as in [74], the return from an asset portfolio is usually represented as a linear function of the portfolio. This is why the standard deviation of the return results in a convex quadratic function. Meanwhile, the profit function in the newsvendor problem is not linear in the order quantities, so that minimizing the standard deviation of the profit ends up as a nonconvex optimization. On the contrary, the CVaR minimization problem has convex structure as long as the profit function is concave or, equivalently, the cost function is convex. Many researchers introduce the standard deviation of the profit in order to capture the profit variation (e.g., [50]) and develop a CAPM by following the modern portfolio theory (e.g., [2]).

In this thesis, we first show that the downside risk measure preserves the convexity, and the resulting risk minimization problem becomes a convex program. We introduce two CVaR measures and achieve several analytical results by making use of this nice structure. More specifically, for the problem of a single product with no constraints, a closed form solution or a simple numerical solution method is derived. We also demonstrate interesting properties of the solutions. We show that the problem is reformulated as a linear program when multiple products under constraints are dealt with and the demand distribution is given as a finite number of scenarios.

1.2.2 α-Conservative Approximation for Probabilistically Constrained Convex Programs

A constraint is said to be a *probabilistic constraint* when it requires a condition be met with a certain probability. Since Charnes, Cooper and Symonds [20] introduced a model involving probabilistic constraints, enormous number of variations have been studied (e.g., [57, 71]). Among them all, convex minimization with a probabilistic constraint, which is called *probabilistically constrained convex program* (PCCP) is the most important problem. The methods proposed to solve general PCCP problems can be roughly categorized into three types:

(a) nonlinear programming methods,

(b) scenario approximation based on Monte Carlo sampling techniques, and

(c) conservative approximation (see Section 2.2 for detailed explanation).

Type (a) is a numerical method using nonlinear optimization techniques, and has been intensively studied for the so-called "random right-hand side" problems (see [27] and references therein). Convexity of the problem is necessary for most of the methods to work, and the gradient of the constraint function is indispensable for some of the methods. Consequently, the methods of this type are not a favorite to be chosen in solving general PCCP.

Type (b) collects a finite number of realizations of the random variables in the model, and builds a convex approximation of the problem by replacing the probabilistic constraint with convex constraints, each of which corresponds to a realization of the random variables. See for example Calafiore and Campi [16, 17].

Type (c) builds a tractable approximation problem by replacing the feasible set of PCCP by a smaller set contained in it. In this sense it is a conservative approach. Among them are Ben-Tal and Nemirovski [7], Bertsimas and Sim [8], Nemirovski [62], and Nemirovski and Shapiro [63]. This conservative approach often faces a criticism that the solution is excessively conservative, which is the drawback shared by type (b) when a large number of realizations are used. This is illustrated by a numerical example given in Nemirovski and Shapiro [63] where the obtained solution satisfies the condition in probability 99% while it is required to satisfy in 95%.

One of the most well-known and practically important examples of the PCCP is the minimization of the Value-at-Risk (VaR) of a financial portfolio. In spite of the several theoretical drawbacks of VaR as a risk measure (e.g., [3]), much attentions still have been paid to the VaR minimization. Though the VaR minimization results in a convex optimization, known as the second-order cone optimization, when the underlying loss follows a normal distribution, it is in general a nonconvex optimization. Typically, the VaR minimization can be formulated as a 0-1 mixed integer program (MIP) when the underlying loss distribution is given by a finite number of scenarios.

Various approaches have been proposed for the VaR minimization. Firstly, heuristic procedures have been proposed in several papers such as the threshold accepting heuristic by Gilli and Këllezi [35], the SP-A model by Puelz [72], the smoothing method by Gaivoronski and Pflug [34], the continuation method by Verma [82], and the CVaR minimization-based algorithms by Larsen, Mausser and Uryasev [48]. Secondly, computationally tractable approximation based on robust optimization techniques has been studied by El Ghaoui, Oks and Oustry [29], and Natarajan, Pachamanova and Sim [61]. Finally, deterministic global optimization approaches have also been explored. For example, Pang and Leyffer [66] formulate this problem as a linear program with linear complementarity constraints and develop a branch-and-bound procedure, and Benati and Rizzi [5] use a mixed integer linear programming formulation for the mean-VaR portfolio program. Furthermore, various studies on the VaR are collected on the web site GloriaMundi.org [36].

Motivated by [63], we consider a conservative approximation of the PCCP with a conservativeness parameter (or approximation accuracy), and apply it to the VaR minimization problem. The resulting problem has a nonconvex feasible region defined by the difference of two convex constraint functions. This formulation is known as the D.C. formulation, and several global or local solution algorithms have been developed (see Tuy [81], for example). A nice feature of the proposed D.C. formulation is that the complexity of the problem is almost independent of the number of scenarios, which contrasts with the fact that the typical MIP formulation requires 0-1 variables as many as the scenarios. We propose a branch-and-bound algorithm, and report some comparative computational results, presenting the performance and characteristics of the proposed algorithm.

1.2.3 Constant Rebalanced Portfolio Optimization under Nonlinear Transaction Costs

Multi-Period Portfolio Optimization. Since the seminal work of Markowitz [53], optimization methods for portfolio selection have been actively studied and are playing an important role in financial decision makings (see, e.g., [22]). Since the early stage, the importance of the multi-period model has been recognized for long-term portfolio management (e.g., [54]). Multi-period portfolio optimization was first introduced as a stochastic control problem (e.g., [55, 75, 56] and [40] for detailed references). Closed-form solutions to those problems need very strong assumptions, and cannot be generalized in the presence of market frictions such as a transaction cost. Moreover, when it comes to a numerical implementation, heavy computation burden should be resolved [15, 14]. Therefore, alternative stochastic programming models have been proposed for multi-period portfolio optimization (e.g., [47, 58, 25, 19]). Mainly studied models, e.g., in [59, 85], employ a scenario tree for representing the uncertainty of asset values. However, to improve the discretization accuracy, the size of the optimization problem grows exponentially as pointed by Ermoliev and Wets [30]. On the other hand, simulated path model, in which scenarios are represented by sample paths generated by a Monte Carlo simulation method, yields a better accuracy in describing uncertainty [38]. Combining the advantages of the two approaches, Hibiki [38] proposed the hybrid model, which is designed not only to describe uncertainty on a simulated path structure but also to enable one to make conditional decisions in a tree structure. In this thesis, taking the compatibility with constant rebalancing strategy into account, the simulated path model is adopted.

Constant Rebalancing Strategy Among investment strategies for the long-term asset management, two extreme policies are buy-and-hold and periodical rebalancing (see [68] for detailed discussion on various investment strategies). Above all else *constant rebalancing* (in other words, fixed mix, constant mix and the like) is the most popular in the latter. This rebalancing strategy requires purchase and sale of assets at the beginning of each period so that investment proportion would be restored to the original one; and the constant rebalancing is a kind of contrarian investment strategy, which suggests to purchase assets with declining price and to sell ones with rising price.

Solution to Constant Rebalanced Portfolio Optimization Multi-period portfolio optimization with constant rebalancing strategy is relatively easy to solve in case of log-optimal portfolio [23], in which the asymptotically optimal portfolio is determined by maximizing the expected log return. However, it becomes a non-convex problem and difficult to attain a globally optimal solution in general when a risk measure (e.g., variance of returns) is introduced [52]. In addition, even any locally optimal solutions of large problems may not be reached via state-of-the-art nonlinear programming solvers. Maranas et al. [52] considered multi-period mean-variance portfolio optimization with constant rebalancing strategy for long-term financial planning. They proposed a rectangular branch-and-bound algorithm in order to globally solve this problem. By enjoying the fact that the number of assets is only up to nine, their deterministic algorithm attains a globally optimal solution in practical time. However, they do not consider the transaction costs and cannot easily deal with it in their framework because introducing cost functions would prevent the problem from having the compact representation they enjoyed. Hibiki [39] proposed an iterative optimization algorithm by alternately fixing decision variables for approximately solving the hybrid model with a fixed-proportion strategy. In this strategy, investment proportions have the same value for all simulated paths passing the same bundle of states. Although this algorithm works well by starting with a good initial solution, it may not work in the constant rebalanced portfolio optimization due to the excessive reduction of degree of freedom in the problem. Besides, transaction costs are not considered in [39].

Transaction Costs. Transaction costs have also been a subject of concerns in the study on portfolio selection, and should be taken into consideration for successful investing in practice. In particular, a large amount of transactions usually make asset price move in an unfavorable direction. Such an effect, known as market impact, makes nonnegligible cost that institutional investors should incur. In this thesis, we consider a convex cost function representing market impact. Among recent papers on multi-period portfolio optimization under transaction costs are the model minimizing one-sided deviation measure [70], the robust optimization approach [11] and the policy optimization approach [18, 78]. However in those papers, constant rebalancing strategy is not considered, and only linear transaction costs are considered. Gaivoronski and Stella [34] proposed a log-optimal portfolio with transaction costs for an adaptive portfolio selection policy. Zhang and Zhang [84] proposed a hybrid model under linear transaction costs in which CVaR is employed as a risk measure, and solved the resulting nonconvex program by applying a genetic algorithm.

We propose a local search algorithm for solving the constant rebalanced portfolio optimization under nonlinear transaction costs. In this algorithm, linear approximation problems and nonlinear equations are iteratively solved via linear programming solver and Newton's method, respectively. In contrast to the use of nonlinear programming solver, the proposed algorithm can provide a solution to the large problem. Moreover, an incumbent solution can be improved better than the alternating optimization [39]. The effectiveness of the proposed local search algorithm is examined through computational experiments where the performance is compared to the buy-and-hold strategy.

1.3 Organization of the Thesis

We discuss the stochastic programming in the next chapter, showing some typical modeling techniques for treating uncertain quantities. We introduce probabilistically constrained convex program. Properties of risk measures VaR and CVaR are also explained. In Chapter 3, CVaR is introduced to the classical newsvendor problem, and closed form solutions of the unconstrained minimization of the two different CVaRs are given. Furthermore, the case of multiple products is analyzed, and when their demand follows a distribution on discrete supports, linear programming formulations are presented. We show differences among the models through several numerical experiments, and we see that the linear programs can be solved efficiently on a personal computer. In Chapter 4, a new conservative approximation for probabilistically constrained convex program is proposed. We propose a branch-and-bound algorithm for solving the new approximation problem of the VaR minimization and present computational results, showing the comparative superiority of the proposed approach. In Chapter 5, a mathematical description of a constant rebalancing model under transaction costs is given, and the portfolio optimization problems via CVaR are explained. We propose the local search algorithm for solving them, and present computational results, showing the comparative superiority of the proposed approach and the constant rebalancing strategy. Several concluding remarks are given in the last chapter.

Chapter 2

Stochastic Programming

A *general form of optimization problem* is formulated as follows:

$$\left| \begin{array}{ll} \underset{\boldsymbol{x} \in \mathbb{R}^J}{\text{minimize}} & \hat{f}(\boldsymbol{x}) \\ \text{subject to} & \hat{g}_i(\boldsymbol{x}) \leq 0, \quad i \in \mathcal{I} := \{1, ..., I\} \\ & \boldsymbol{x} \in X, \end{array} \right. \tag{2.1}$$

where the function $\hat{f} : \mathbb{R}^J \to \mathbb{R}$ is called the *objective function*, and the set

$$\{\boldsymbol{x} \in \mathbb{R}^J \mid \hat{g}_i(\boldsymbol{x}) \leq 0, \ i \in \mathcal{I}; \quad \boldsymbol{x} \in X\},$$

is called the *feasible region*. In many real-life applications, there might be an uncertainty about parameters involved in the functions $\hat{f}$ and $\hat{g}_i$. In this case, we consider the alternative optimization problem by introducing random variables. Let $(\Omega, \mathcal{F})$ be a sample space, equipped with the sigma algebra $\mathcal{F}$. We denote by $\tilde{\boldsymbol{\xi}} = \boldsymbol{\xi}(\omega)$ a d-dimensional real random vector and by Ξ ($\subseteq \mathbb{R}^d$) its support. We have the following optimization problem under uncertain environment:

$$\left| \begin{array}{ll} \underset{\boldsymbol{x} \in \mathbb{R}^J}{\text{minimize}} & f(\boldsymbol{x}, \tilde{\boldsymbol{\xi}}) \\ \text{subject to} & g_i(\boldsymbol{x}, \tilde{\boldsymbol{\xi}}) \leq 0, \quad i \in \mathcal{I} \\ & \boldsymbol{x} \in X, \end{array} \right. \tag{2.2}$$

where the functions $f : \mathbb{R}^J \times \Xi \to \mathbb{R}$ and $g_i : \mathbb{R}^J \times \Xi \to \mathbb{R}$ model random outcomes. However, Problem (2.2) is not clearly defined before knowing the realization of $\tilde{\boldsymbol{\xi}}$ since the meanings of the objective as well as of the constraints are ambiguous. Therefore, the functions $\hat{f}$ and $\hat{g}_i$ need to be defined as $\hat{f}(\boldsymbol{x}) := F[f(\boldsymbol{x}, \tilde{\boldsymbol{\xi}})]$ and

$\hat{g}_i(\boldsymbol{x}) := G_i[g_i(\boldsymbol{x}, \tilde{\boldsymbol{\xi}})]$ so that the uncertain optimization problem (2.2) would be reduced to the form of deterministic optimization problem (2.1). A common recipe is to use the expectation as F and G_i:

$$\left|\begin{array}{ll} \underset{\boldsymbol{x}\in\mathbb{R}^J}{\text{minimize}} & \mathbb{E}[f(\boldsymbol{x}, \tilde{\boldsymbol{\xi}})] \\ \text{subject to} & \mathbb{E}[g_i(\boldsymbol{x}, \tilde{\boldsymbol{\xi}})] \leq 0, \quad i \in \mathcal{I} \\ & \boldsymbol{x} \in X. \end{array}\right. \tag{2.3}$$

As far as the long-term performance is concerned, the expected value of the objective function is appropriate. However, this approach disregards the risk of incurring a large loss. In Section 2.4, we show risk-averse optimization approaches by adopting several risk measures. In addition, we should notice that a feasible solution of Problem (2.3) may not satisfy the original constraint $\hat{g}_i(\boldsymbol{x}, \tilde{\boldsymbol{\xi}}) \leq 0,\ i \in \mathcal{I}$ with high probability. Feasible approaches to the matter are Robust Optimization (see Section 2.1) which forces the constraints to be satisfied for any realizations of $\tilde{\boldsymbol{\xi}}$, and Probabilistic Constraint (see Section 2.2) which compels a solution to satisfy the constraints with certain probability. In Section 2.3, we also explain a Two-Stage Problem where the penalty for not satisfying the constraints is added to the objective function.

2.1 Robust Optimization

One remedy for the constraints including random variables in Problem (2.2) is a robust optimization approach where the constraint functions are defined as $\sup\{g_i(\boldsymbol{x}, \boldsymbol{\xi}) \mid \boldsymbol{\xi} \in \Xi\}$ for $i \in \mathcal{I}$.

Robust optimization problem is

$$\left|\begin{array}{ll} \underset{\boldsymbol{x}\in\mathbb{R}^J}{\text{minimize}} & \hat{f}(\boldsymbol{x}) \\ \text{subject to} & \sup\{g_i(\boldsymbol{x}, \boldsymbol{\xi}) \mid \boldsymbol{\xi} \in \Xi\} \leq 0, \quad i \in \mathcal{I} \\ & \boldsymbol{x} \in X, \end{array}\right. \tag{2.4}$$

or equivalently:

$$\left|\begin{array}{ll} \underset{\boldsymbol{x}\in\mathbb{R}^J}{\text{minimize}} & \hat{f}(\boldsymbol{x}) \\ \text{subject to} & g_i(\boldsymbol{x}, \boldsymbol{\xi}) \leq 0, \quad \boldsymbol{\xi} \in \Xi, \quad i \in \mathcal{I} \\ & \boldsymbol{x} \in X. \end{array}\right. \tag{2.5}$$

Namely, a feasible solution of the problem should satisfy $g_i(\boldsymbol{x}, \tilde{\boldsymbol{\xi}}) \leq 0,\ i \in \mathcal{I}$ for any realizations of $\tilde{\boldsymbol{\xi}}$. We assume without loss of generality that the objective function is not affected by uncertainty since the robust counterpart of Problem (2.2), which is the minimization of the supremum of the objective value over Ξ, can be formulated as:

$$\left|\begin{array}{ll} \underset{(\boldsymbol{x},t)\in \mathbb{R}^J \times \mathbb{R}}{\text{minimize}} & t \\ \text{subject to} & \sup\{f(\boldsymbol{x}, \boldsymbol{\xi}) - t \mid \boldsymbol{\xi} \in \Xi\} \leq 0 \\ & \sup\{g_i(\boldsymbol{x}, \boldsymbol{\xi}) \mid \boldsymbol{\xi} \in \Xi\} \leq 0, \qquad i \in \mathcal{I} \\ & \boldsymbol{x} \in X. \end{array}\right. \tag{2.6}$$

When the functions g_i as well as $\hat{f}$ are convex in $\boldsymbol{x}$ for any fixed $\boldsymbol{\xi} \in \Xi$, the robust optimization problem is a convex minimization. However, the robust optimization problem is intractable in general due to infinitely many constraints.

To present a framework of the robust optimization, we consider the following linear programming problem including random parameters $\tilde{\boldsymbol{a}}_i$:

$$\left|\begin{array}{ll} \underset{\boldsymbol{x}\in \mathbb{R}^J}{\text{minimize}} & \boldsymbol{c}^\top \boldsymbol{x} \\ \text{subject to} & \tilde{\boldsymbol{a}}_i^\top \boldsymbol{x} \geq b_i, \quad i \in \mathcal{I} \\ & \boldsymbol{x} \in X, \end{array}\right. \tag{2.7}$$

where X is a polyhedron. We assume that any realization of the matrix $\tilde{\boldsymbol{a}}_i$ belongs to the set $\mathcal{A}_i$ and refer to $\mathcal{A}_i$ as the *uncertainty set*.

It is shown in [6] that for any convex uncertainty set, the robust counterpart of Problem (2.7) can be reformulated as a single convex programming problem. For instance in the case of ellipsoidal uncertainty set, the robust counterpart of Problem (2.7) is reduced to a second-order cone problem (e.g., [6]). Although the robust optimization approach has the drawback of computational complexity, there is still a case that the robust counterpart of a linear programming problem remains a linear programming problem (e.g., [8, 10]).

To present the framework proposed by [8], we assume that every coefficient $\tilde{a}_{ij}$ is subject to uncertainty, and its realization belongs to the interval $[\bar{a}_{ij} - \hat{a}_{ij},\ \bar{a}_{ij} + \hat{a}_{ij}]$. We define the scaled deviation z_{ij} of parameter $\tilde{a}_{ij}$ from its nominal value $\bar{a}_{ij}$ as

$$z_{ij} = \frac{\tilde{a}_{ij} - \bar{a}_{ij}}{\hat{a}_{ij}}. \tag{2.8}$$

The scaled deviation of a parameter always belongs to $[-1, 1]$.

Then, the robust counterpart of Problem (2.7) is formulated as

$$\left|\begin{array}{ll} \underset{\boldsymbol{x}\in\mathbb{R}^J}{\text{minimize}} & \boldsymbol{c}^\top \boldsymbol{x} \\ \text{subject to} & \bar{\boldsymbol{a}}_i^\top \boldsymbol{x} + \min\left\{ \sum_{j\in\mathcal{J}} \hat{a}_{ij} x_j z_{ij} \;\middle|\; \boldsymbol{z}_i \in \mathcal{Z}_i \right\} \geq b_i, \quad i \in \mathcal{I} \\ & \boldsymbol{x} \in X, \end{array}\right. \tag{2.9}$$

where $\boldsymbol{z}_i$ is the vector whose j-th element is z_{ij}, and $\mathcal{Z}_i$ is defined as

$$\mathcal{Z}_i = \left\{ \boldsymbol{z} \in \mathbb{R}^J \;\middle|\; |z_j| \leq 1,\ j \in \mathcal{J} = \{1, ..., J\};\ \ \sum_{j\in\mathcal{J}} |z_j| \leq \Gamma_i \right\}. \tag{2.10}$$

The parameter $\Gamma_i \in [0, J]$ represents the degree of conservatism of the solution. The degree of conservatism is mitigated when $\Gamma_i < J$, and the uncertain coefficients $\bar{\boldsymbol{a}}_i$ would be constant vectors if $\Gamma_i = 0$.

Noting that $\hat{a}_{ij} \geq 0$, one has

$$\begin{aligned} &\min\left\{ \sum_{j\in\mathcal{J}} \hat{a}_{ij} x_j z_{ij} \;\middle|\; \boldsymbol{z}_i \in \mathcal{Z}_i \right\} \\ &= -\max\left\{ \sum_{j\in\mathcal{J}} \hat{a}_{ij} |x_j| z_{ij} \;\middle|\; \sum_{j\in\mathcal{J}} z_{ij} \leq \Gamma_i;\ \ 0 \leq z_{ij} \leq 1,\ j \in \mathcal{J} \right\}. \end{aligned} \tag{2.11}$$

Applying the strong duality argument (see [8] for details), we then reformulate Problem (2.9) as the following linear programming problem:

$$\left|\begin{array}{lll} \underset{\boldsymbol{x},\boldsymbol{y},\boldsymbol{u},\boldsymbol{v}}{\text{minimize}} & \boldsymbol{c}^\top \boldsymbol{x} & \\ \text{subject to} & \bar{\boldsymbol{a}}_i^\top \boldsymbol{x} - \Gamma_i u_i - \sum_{j=1}^{n} v_{ij} \geq b_i, & i \in \mathcal{I} \\ & u_i + v_{ij} \geq \hat{a}_{ij} y_j, & i \in \mathcal{I},\ j \in \mathcal{J} \\ & -y_j \leq x_j \leq y_j, & j \in \mathcal{J} \\ & u_i,\ v_{ij} \geq 0, & i \in \mathcal{I},\ j \in \mathcal{J} \\ & \boldsymbol{x} \in X. & \end{array}\right. \tag{2.12}$$

2.2 Probabilistic Constraint

In the robust optimization, the solution must satisfy the constraints for any realization of random variables, and it is often too conservative for practical implementation. Compared to the robust optimization, probabilistic constraint (also called chance constraint) is more flexible since it forces the constraints to be satisfied with certain probability.

Let $\beta \in (0,1)$ and $\beta_i \in (0,1)$, $i \in \mathcal{I} = \{1, ..., I\}$ be parameters representing the *confidence level*. In the case where $I \geq 2$, the constraint

$$\mathrm{Prob}\{g_i(\boldsymbol{x}, \tilde{\boldsymbol{\xi}}) \leq 0\} \geq \beta_i, \tag{2.13}$$

is referred to as the *individual probabilistic constraint*, while the constraint

$$\mathrm{Prob}\{g_i(\boldsymbol{x}, \tilde{\boldsymbol{\xi}}) \leq 0, \ {}^\forall i \in \mathcal{I}\} \geq \beta, \tag{2.14}$$

is referred to as the *jointly probabilistic constraint*. We here consider the jointly probabilistic constraint.

Assuming that the objective function includes no uncertainty, *probabilistically constrained problem* is formulated as follows:

$$\left| \begin{array}{ll} \underset{\boldsymbol{x} \in \mathrm{I\!R}^J}{\text{minimize}} & \hat{f}(\boldsymbol{x}) \\ \text{subject to} & \mathrm{Prob}\{g_i(\boldsymbol{x}, \tilde{\boldsymbol{\xi}}) \leq 0, \ {}^\forall i \in \mathcal{I}\} \geq \beta \\ & \boldsymbol{x} \in X. \end{array} \right. \tag{2.15}$$

The probabilistic constraint almost coincides with the robust constraint when β is close to 1. In spite of this, the situation becomes more difficult for probabilistically constrained problems as demonstrated below. A point $\bar{\boldsymbol{x}} \in \mathrm{I\!R}^J$ is a feasible solution to Problem (2.15) iff we find at least one event $G \in \mathcal{F}$ such that $\mathrm{Prob}\, G \geq \beta$ and $g_i(\bar{\boldsymbol{x}}, \boldsymbol{\xi}(\omega)) \leq 0$ for all $\omega \in G$ and $i \in \mathcal{I}$. Therefore, if $\mathcal{G} \subseteq \mathcal{F}$ is the collection of all events G of $\mathcal{F}$ such that $\mathrm{Prob}\, G \geq \beta$, then the feasible set described by the probabilistic constraint can be rewritten as

$$\begin{aligned} &\{\boldsymbol{x} \in \mathrm{I\!R}^J \mid \mathrm{Prob}\{g_i(\boldsymbol{x}, \tilde{\boldsymbol{\xi}}) \leq 0, \ {}^\forall i \in \mathcal{I}\} \geq \beta\} \\ &= \bigcup_{G \in \mathcal{G}} \bigcap_{\omega \in G} \{\boldsymbol{x} \in \mathrm{I\!R}^J \mid g_i(\boldsymbol{x}, \boldsymbol{\xi}(\omega)) \leq 0, \ {}^\forall i \in \mathcal{I}\}. \end{aligned} \tag{2.16}$$

Since the union of convex sets need not be convex, (2.16) implies that the feasible region of Problem (2.15) may not be convex even if $\{\boldsymbol{x} \in \mathrm{I\!R}^J \mid g_i(\boldsymbol{x}, \boldsymbol{\xi}(\omega)) \leq 0, \ {}^\forall i \in \mathcal{I}\}$ is convex for every $\omega \in \Omega$. This is an undesirable property of the probabilistic constraint compared with the robust constraint.

The following results are well-known.

Proposition 2.1 (Section 1.5, Kall and Wallace [41])
Let the probability measure Prob *have a density function* ϕ*. Then* $\{\boldsymbol{x} \in \mathrm{IR}^J \mid \mathrm{Prob}\{g_i(\boldsymbol{x}, \tilde{\boldsymbol{\xi}}) \leq 0, \ ^\forall i \in \mathcal{I}\} \geq \beta\}$ *is convex if*

- $\log \phi$ *is a concave function or* $\phi^{-1/d}$ *is a quasi-concave function,*
- $g_i(\boldsymbol{x}, \boldsymbol{\xi})$ *is convex jointly in both arguments* $\boldsymbol{x}$ *and* $\boldsymbol{\xi}$ *for all* $i \in \mathcal{I}$*.* $\star$

Proposition 2.2 (Section 1.5, Kall and Wallace [41])
Let the probability measure Prob *have a density function* ϕ*. Then the feasible region described by random right-hand-side probabilistic constraint* $\{\boldsymbol{x} \in \mathrm{IR}^J \mid \mathrm{Prob}\{\boldsymbol{a}_i^\top \boldsymbol{x} \geq \tilde{b}_i, \ ^\forall i \in \mathcal{I}\} \geq \beta\}$ *is convex if* $\log \phi$ *is a concave function or* $\phi^{-1/d}$ *is a quasi-concave function.* $\star$

2.2.1 Solution to Random Right-Hand-Side Problem

Despite various computational methods proposed in the literatures, the general probabilistically constrained problem is difficult to solve efficiently. One of the types of efficiently solved problems is the so-called *random right-hand-side probabilistically constrained linear programming problem*:

$$
\left| \begin{array}{ll} \underset{\boldsymbol{x} \in \mathrm{IR}^J}{\text{minimize}} & \hat{f}(\boldsymbol{x}) \\ \text{subject to} & \mathrm{Prob}\{\boldsymbol{a}_i^\top \boldsymbol{x} \geq \tilde{b}_i, \ ^\forall i \in \mathcal{I}\} \geq \beta \\ & \boldsymbol{x} \in X, \end{array} \right. \tag{2.17}
$$

where $\boldsymbol{a}_i$ are constant vectors, and $\tilde{b}_i$ are uncertain parameters and X is a polytope. In the following, $H(\boldsymbol{x})$ is defined as $H(\boldsymbol{x}) := \mathrm{Prob}\{\boldsymbol{a}_i^\top \boldsymbol{x} \geq \tilde{b}_i, \ ^\forall i \in \mathcal{I}\}$.

Numerical techniques especially for solving the random right-hand-side probabilistically constrained linear programming problem (2.17) have been developed. We review two of the existing methods based on nonlinear programming methods by following Dentcheva [27].

Cutting Plane Method One of the methods is based on the cutting plane technique. It is assumed that

- the objective function is linear, i.e., $\hat{f}(\boldsymbol{x}) := \boldsymbol{c}^\top \boldsymbol{x}$ where $\boldsymbol{c}$ is a constant vector,

- the constraint function H is quasi-concave,
- the constraint function H has a continuous gradient,
- the following Slater constraint qualification condition is satisfied: there exists a vector $\boldsymbol{x}^0 \in X$ such that

$$H(\boldsymbol{x}^0) > \beta. \tag{2.18}$$

The underlying principle of the algorithm is to solve a sequence of approximating linear programming problems and to approximate the feasible region by cutting off the current solution. Under the above assumptions, algorithm steps are as follows.

Cutting Plane Method for Problem (2.17)

Step 0. [initialization.] Set $P^1 \leftarrow X$ and $k \leftarrow 1$.

Step 1. [Solution of LP problem.] Solve the linear programming problem:

$$\left| \underset{\boldsymbol{x} \in \mathbb{R}^J}{\text{minimize}} \ \boldsymbol{c}^\top \boldsymbol{x} \ \ \text{subject to } \boldsymbol{x} \in P^k. \right.$$

Let $\boldsymbol{x}^k$ be an optimal solution. If $\boldsymbol{x}^k$ is a feasible solution to the original problem, then stop, and $\boldsymbol{x}^k$ is an optimal solution. Otherwise, go to Step 2.

Step 2. [Addition of the cutting plane.] Let λ^k be the largest $\lambda \geq 0$ such that $\boldsymbol{x}^0 + \lambda(\boldsymbol{x}^k - \boldsymbol{x}^0)$ is feasible and set

$$\boldsymbol{y}^k := \boldsymbol{x}^0 + \lambda^k(\boldsymbol{x}^k - \boldsymbol{x}^0).$$

Then define

$$P^{k+1} := \{\boldsymbol{x} \in \mathbb{R}^J \mid \boldsymbol{x} \in P^k, \ \nabla H(\boldsymbol{y}^k)(\boldsymbol{x} - \boldsymbol{y}^k) \geq 0\}.$$

Set $k \leftarrow k + 1$ and go to Step 1.

Logarithmic Barrier Function Method If the logarithm of the constraint function, i.e., $\log H(\boldsymbol{x})$ is concave, then we can solve Problem (2.17) by utilizing logarithmic penalty functions. Let $\{\ell^k\}$ be a decreasing sequence of positive numbers such that $\lim_{k\to\infty} \ell^k = 0$ and consider the following convex problem:

$$\left|\begin{array}{ll}\underset{\boldsymbol{x}\in\mathbb{R}^J}{\text{minimize}} & \hat{f}(\boldsymbol{x}) - \ell^k \log(H(\boldsymbol{x}) - \beta) \\ \text{subject to} & \boldsymbol{x} \in X.\end{array}\right. \tag{2.19}$$

We solve a sequence of Problems (2.19) in this method. Let $\boldsymbol{x}^k$ be an optimal solution to Problem (2.19). Then the sequence $\{f(\boldsymbol{x}^k)\}$ converges to the optimal value of Problem (2.17) as $k \to \infty$ under the following assumptions:

- the objective function $\hat{f}$ is continuous and convex.
- the constraint function H is continuous.
- the logarithm of the constraint function, i.e., $\log H$ is concave.
- the Slater constraint qualification condition (2.18) is satisfied.

2.2.2 Solution to Probabilistically Constrained Convex Program

In this thesis, we focus on a *probabilistically constrained convex program* (PCCP). We assume in the problem that

- the function $\hat{f}(\boldsymbol{x})$ is convex in $\boldsymbol{x}$,
- the set X is closed convex, and
- the scalar-valued function $g(\boldsymbol{x}, \boldsymbol{\xi})$ is convex in $\boldsymbol{x}$ for each $\boldsymbol{\xi} \in \Xi$.

PCCP is formulated as follows:

$$\text{(PCCP)}\ \left|\begin{array}{ll}\underset{\boldsymbol{x}\in\mathbb{R}^J}{\text{minimize}} & \hat{f}(\boldsymbol{x}) \\ \text{subject to} & \text{Prob}\Big\{g(\boldsymbol{x}, \tilde{\boldsymbol{\xi}}) > 0\Big\} \leq 1 - \beta \\ & \boldsymbol{x} \in X,\end{array}\right. \tag{2.20}$$

and the probability

$$\text{VP}(\boldsymbol{x}) := \text{Prob}\Big\{g(\boldsymbol{x}, \tilde{\boldsymbol{\xi}}) > 0\Big\}, \tag{2.21}$$

is called the *violation probability*. The function g is assumed to be scalar-valued without loss of generality. Indeed, the jointly probabilistic constraint (2.14) where the functions $g_i(\boldsymbol{x}, \boldsymbol{\xi})$ are convex in $\boldsymbol{x}$ for any fixed $\boldsymbol{\xi}$ can be converted into the constraint in PCCP (2.20) by putting $g(\boldsymbol{x}, \boldsymbol{\xi}) := \max\{g_i(\boldsymbol{x}, \boldsymbol{\xi}) \mid i \in \mathcal{I}\}$.

Despite the convexity of functions $\hat{f}$ and g, this problem has the nonconvex feasible region in general, and consequently, is intractable. In particular, it can have multiple local minima when the support of the associated probabilities is given by finitely many scenarios.

Scenario Approximation We explain a *scenario approximation method* for solving PCCP (2.20). Let us assume that we have a sample $\{\, \boldsymbol{\xi}_s \mid s \in \mathcal{S} = \{1, ..., S\}\}$ of S realizations of the random vector $\tilde{\boldsymbol{\xi}}$. Then, we consider the following problem:

$$\left|\begin{array}{ll} \underset{\boldsymbol{x} \in \mathbb{R}^J}{\text{minimize}} & \hat{f}(\boldsymbol{x}) \\ \text{subject to} & g(\boldsymbol{x}, \boldsymbol{\xi}_s) \leq 0, \;\; s = 1, ..., S \\ & \boldsymbol{x} \in X. \end{array}\right. \tag{2.22}$$

When the confidence level β is close to 1, Problem (2.22) can be regarded as an approximation of Problem (2.15). Furthermore if the number of constraints S is very large, it is expected to see that a feasible solution of Problem (2.22) becomes a feasible solution of Problem (2.20) with high probability. Moreover, the feasible region of Problem (2.22) is convex since the intersection of convex sets is convex. However, another crucial problem would be how to determine the number S, i.e., how many scenarios are needed in order to guarantee that the solution of Problem (2.22) satisfies the probabilistic constraint of Problem (2.20). Calafiore and Campi [16, 17] showed the following results on a feasible solution $\boldsymbol{x}_S$ of Problem (2.22):

Theorem 2.3 (Calafiore and Campi [17])

$$\mathbb{E}\Big[\mathrm{Prob}\{g(\boldsymbol{x}_S, \tilde{\boldsymbol{\xi}}) \leq 0\}\Big] \geq 1 - \frac{J}{S+1}.$$

⋆

Theorem 2.4 (Calafiore and Campi [16])
For $\varepsilon \in (0, 1)$*, if*

$$S \geq \frac{2}{1-\beta} \log \frac{1}{\varepsilon} + 2J + \frac{2J}{1-\beta} \log \frac{2}{1-\beta},$$

then $\mathrm{Prob}\{g(\boldsymbol{x}_S, \tilde{\boldsymbol{\xi}}) \leq 0\} \geq \beta$ *with probability higher than* $1 - \varepsilon$. ⋆

The intrinsic drawback of this type of approximation is that the obtained solution is too conservative in terms of the violation probability VP (2.21), that is, the violation probability of the obtained solution can be much smaller than $1 - \beta$. Furthermore Problem (2.22) has many nonlinear constraints, which makes it difficult to solve.

Convex Conservative Approximation In Nemirovski and Shapiro [63], a convex conservative constraint is adopted in place of the probabilistic constraint of Problem (2.20). We here describe their approach for PCCP (2.20).

Let $\psi : \mathbb{R} \to \mathbb{R}$ be any nonnegative valued, nondecreasing function satisfying $\psi(0) \geq 1$. For a scalar valued random variable $\tilde{Z}$, one then has

$$\mathbb{E}\Big[\psi(\tilde{Z})\Big] \;\geq\; \mathbb{E}\Big[1\!\!1_{[0,+\infty)}(\tilde{Z})\Big] \;=\; \mathrm{Prob}\{\tilde{Z} \geq 0\} \;\geq\; \mathrm{Prob}\{\tilde{Z} > 0\}. \tag{2.23}$$

From this relation, by taking $\tilde{Z} = g(\boldsymbol{x}, \tilde{\boldsymbol{\xi}})$, it is clear that

$$\Big\{\, \boldsymbol{x} \in \mathbb{R}^J \,\Big|\, \mathbb{E}\Big[\psi(g(\boldsymbol{x}, \tilde{\boldsymbol{\xi}}))\Big] \leq 1 - \beta \,\Big\} \subseteq \Big\{\, \boldsymbol{x} \in \mathbb{R}^J \,\Big|\, \mathrm{VP}(\boldsymbol{x}) \leq 1 - \beta \,\Big\}.$$

By restricting ψ to a convex function, $\psi(g(\boldsymbol{x}, \boldsymbol{\xi}))$ is convex in $\boldsymbol{x}$ for any fixed $\boldsymbol{\xi}$, and $\mathbb{E}\Big[\psi(g(\boldsymbol{x}, \tilde{\boldsymbol{\xi}}))\Big]$ is convex in $\boldsymbol{x}$ due to the property of expectation. Consequently, by replacing the probabilistic constraint in Problem (2.20) with the constraint $\mathbb{E}\Big[\psi(g(\boldsymbol{x}, \tilde{\boldsymbol{\xi}}))\Big] \leq 1 - \beta$, we obtain a convex conservative approximation problem as

$$\text{(CCAP)} \quad \left|\; \begin{array}{ll} \underset{\boldsymbol{x} \in \mathbb{R}^J}{\text{minimize}} & \hat{f}(\boldsymbol{x}) \\ \text{subject to} & \mathbb{E}\Big[\psi(g(\boldsymbol{x}, \tilde{\boldsymbol{\xi}}))\Big] \leq 1 - \beta \\ & \boldsymbol{x} \in X. \end{array} \right.$$

It should be noted that the feasible region is always contained in that of Problem (2.20), and accordingly, any feasible solution to this problem is feasible to Problem (2.20).

Nemirovski and Shapiro [63] propose a tighter convex conservative approximation as follows. From the relation (2.23) we also have

$$\mathbb{E}\Big[\psi\Big(\frac{\tilde{Z}}{t}\Big)\Big] \;\geq\; \mathrm{Prob}\Big\{\frac{\tilde{Z}}{t} > 0\Big\} \;=\; \mathrm{Prob}\Big\{\tilde{Z} > 0\Big\},$$

for any positive scalar value t. Similarly, by taking $\tilde{Z} = g(\boldsymbol{x}, \tilde{\boldsymbol{\xi}})$, we see that the following condition is also a conservative constraint:

$$\mathbb{E}\Big[\psi\Big(\frac{1}{t}g(\boldsymbol{x}, \tilde{\boldsymbol{\xi}})\Big)\Big] \leq (1-\beta).$$

By multiplying the above relation by $t > 0$ and by exploiting the fact that a perspective operation preserves convexity (see Section 3.2.6, [13] for instance), we obtain the following convex conservative constraint for PCCP (2.20):

$$\inf\Big\{t\,\mathbb{E}\Big[\psi\Big(\frac{1}{t}g(\boldsymbol{x}, \tilde{\boldsymbol{\xi}})\Big)\Big] - t(1-\beta)\,\Big|\, t > 0\Big\} \leq 0. \tag{2.24}$$

Though the left-hand side of the constraint (2.24) is the infimum over $t > 0$, we solve the resulting conservative approximation problem via an one-level (nonlinear) convex optimization:

$$\left|\begin{array}{ll} \underset{(\boldsymbol{x},t)\in\mathbb{R}^n\times\mathbb{R}}{\text{minimize}} & \hat{f}(\boldsymbol{x}) \\ \text{subject to} & t\,\mathbb{E}\Big[\psi\Big(\frac{1}{t}g(\boldsymbol{x}, \tilde{\boldsymbol{\xi}})\Big)\Big] - t(1-\beta) \leq 0 \\ & t \geq 0 \\ & \boldsymbol{x} \in X. \end{array}\right. \tag{2.25}$$

A criticism of this approach is that it often provides too conservative solutions. In order to overcome this drawback, they propose to replace β iteratively with a smaller value β^- in (2.25) and solve this problem until the violation probability becomes close to $1-\beta$. Although this strategy may succeed in finding a feasible solution to the original problem (2.20) with a higher violation probability, it is highly probable that the obtained objective value is much worse than the optimal value of the original problem (2.20). In Chapter 4, we propose a new conservative constraint and to solve the corresponding conservative approximation problem by applying global optimization algorithms.

2.3 Two-Stage Problem

As we have stated, it is not appropriate to take only the expected value of the constraint function in Problem (2.2) into consideration since the solution to be obtained may not satisfy the original constraints with sufficiently high probability.

In this section, we will give a brief introduction of the two-stage problem by following Birge and Louveaux [12].

The problem is the following linear programming problem including random variables:

$$\left|\begin{array}{ll} \underset{\boldsymbol{x}\in\mathbb{R}^J}{\text{minimize}} & \boldsymbol{c}^\top \boldsymbol{x} \\ \text{subject to} & \boldsymbol{A}\boldsymbol{x} \geq \boldsymbol{b} \\ & \tilde{\boldsymbol{T}}\boldsymbol{x} = \tilde{\boldsymbol{h}} \\ & \boldsymbol{x} \geq \boldsymbol{0}, \end{array}\right. \tag{2.26}$$

where $\boldsymbol{c}$ and $\boldsymbol{b}$ are constant vectors, and $\boldsymbol{A}$ is a constant matrix, and matrix $\tilde{\boldsymbol{T}}$ and vector $\tilde{\boldsymbol{h}}$ are subject to uncertainty. We define the recourse problem as follows:

$$\tilde{Q}(\boldsymbol{x}) := \min_{\boldsymbol{y}}\{\boldsymbol{q}^\top \boldsymbol{y} \mid \boldsymbol{W}\boldsymbol{y} = \tilde{\boldsymbol{h}} - \tilde{\boldsymbol{T}}\boldsymbol{x},\ \boldsymbol{y} \geq \boldsymbol{0}\}, \tag{2.27}$$

where $\boldsymbol{q}$ is a constant vector, and $\boldsymbol{W}$ is a constant matrix, and $\boldsymbol{y}$ is a second-stage decision variable. $\tilde{Q}(\boldsymbol{x})$ is referred to as the recourse cost function which represents the penalty for not satisfying the constraint $\tilde{\boldsymbol{T}}\boldsymbol{x} = \tilde{\boldsymbol{h}}$.

Linear two-stage problem is formulated as follows:

$$\left|\begin{array}{ll} \underset{\boldsymbol{x}\in\mathbb{R}^J}{\text{minimize}} & \boldsymbol{c}^\top \boldsymbol{x} + \mathbb{E}[\tilde{Q}(\boldsymbol{x})] \\ \text{subject to} & \boldsymbol{A}\boldsymbol{x} \geq \boldsymbol{b} \\ & \boldsymbol{x} \geq \boldsymbol{0} \\ & \tilde{Q}(\boldsymbol{x}) = \min_{\boldsymbol{y}}\{\boldsymbol{q}^\top \boldsymbol{y} \mid \boldsymbol{W}\boldsymbol{y} = \tilde{\boldsymbol{h}} - \tilde{\boldsymbol{T}}\boldsymbol{x},\ \boldsymbol{y} \geq \boldsymbol{0}\}, \end{array}\right. \tag{2.28}$$

where the sum of the original objective function $\boldsymbol{c}^\top \boldsymbol{x}$ and the expected value of the recourse cost $\tilde{Q}(\boldsymbol{x})$ is minimized.

Although Problem (2.28) is a sort of two-level optimization problem, Problem (2.28) becomes a convex problem under a weak condition. L-Shaped Method, which is a version of the well-known Bender's decomposition method (see e.g., [12, 41]), is one of the solution methods under the assumption that the uncertain parameters $\tilde{\boldsymbol{T}}$ and $\tilde{\boldsymbol{h}}$ follow discrete distribution. The classic newsvendor problem we will consider in Chapter 3 is a special case of Problem (2.28).

2.4 Risk Averse Optimization via Value-at-Risk and Conditional Value-at-Risk

In this section, we discuss the two kinds of typical risk measures, the value-at-risk (VaR) and the conditional value-at-risk (CVaR), which are widely adopted in stochastic programming problems. We see that the CVaR satisfies various desirable properties as a risk measure.

2.4.1 Properties of VaR and CVaR

Following Rockafellar and Uryasev [74], we introduce the CVaR.

Let $\mathcal{L}(\boldsymbol{x}, \tilde{\boldsymbol{\xi}})$ denote a loss which is a random variable for each $\boldsymbol{x}$, and let us denote the distribution function of $\mathcal{L}$ by $F_{\mathcal{L}}(\eta \,|\, \boldsymbol{x}) := \mathrm{Prob}\left\{\mathcal{L}(\boldsymbol{x}, \tilde{\boldsymbol{\xi}}) \leq \eta\right\}$. Anything can be adopted as the loss $\mathcal{L}$ if we prefer it to be small.

For $\beta \in [0, 1)$, we define the *β-value-at-risk* of the distribution by

$$\mathrm{VaR}_\beta(\boldsymbol{x}) := \min \{\, m \,|\, F_{\mathcal{L}}(m|\boldsymbol{x}) \geq \beta \,\} \,.$$

By definition, the loss $\mathcal{L}$ exceeds VaR_β with only probability lower than $1 - \beta$.

Rockafellar and Uryasev [74] introduce the β-tail distribution function to focus on the upper tail part of the loss distribution as

$$F_{\mathcal{L}}^{\beta}(\eta \,|\, \boldsymbol{x}) := \begin{cases} 0 & \text{for } \eta < \mathrm{VaR}_\beta(\boldsymbol{x}), \\ \dfrac{F_{\mathcal{L}}(\eta \,|\, \boldsymbol{x}) - \beta}{1 - \beta} & \text{for } \eta \geq \mathrm{VaR}_\beta(\boldsymbol{x}). \end{cases}$$

Using the expectation operator $\mathbb{E}_\beta[\cdot]$ under the β-tail distribution $F_{\mathcal{L}}^{\beta}$, we define the *β-conditional value-at-risk* of the loss $\mathcal{L}$ by

$$\mathrm{CVaR}_\beta(\boldsymbol{x}) := \mathbb{E}_\beta \left[\mathcal{L}(\boldsymbol{x}, \tilde{\boldsymbol{\xi}}) \right].$$

Denoting the expectation under the original distribution $F_{\mathcal{L}}$ by $\mathbb{E}[\cdot]$, the following relation is shown in [74]:

$$\mathbb{E}[\, \mathcal{L}(\boldsymbol{x}, \tilde{\boldsymbol{\xi}}) \,|\, \mathcal{L}(\boldsymbol{x}, \tilde{\boldsymbol{\xi}}) \geq \mathrm{VaR}_\beta(\boldsymbol{x}) \,] \;\leq\; \mathrm{CVaR}_\beta(\boldsymbol{x}) \;\leq\; \mathbb{E}[\, \mathcal{L}(\boldsymbol{x}, \tilde{\boldsymbol{\xi}}) \,|\, \mathcal{L}(\boldsymbol{x}, \tilde{\boldsymbol{\xi}}) > \mathrm{VaR}_\beta(\boldsymbol{x}) \,], \tag{2.29}$$

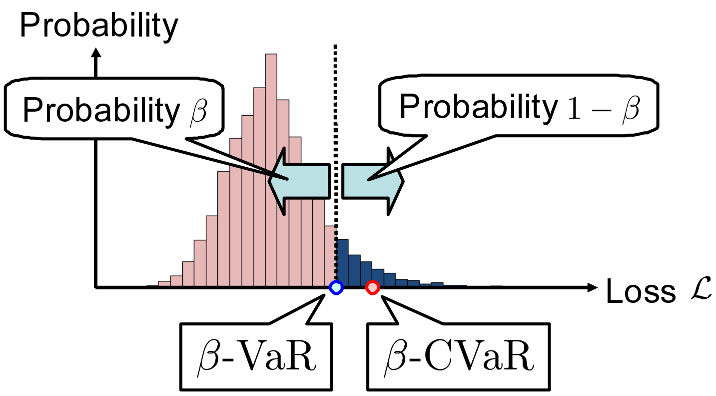


Figure 2.1: Value-at Risk and Conditional Value-at-Risk

which implies that CVaR_β is approximately equal to the conditional expectation of $\mathcal{L}$ which exceeds the threshold $\mathrm{VaR}_\beta(\boldsymbol{x})$ for fixed $\boldsymbol{x}$. See Figure 2.1.

In order to minimize $\mathrm{CVaR}_\beta(\boldsymbol{x})$, Rockafellar and Uryasev [74] introduce an auxiliary function $G_\beta : \mathbb{R}^{J+1} \to \mathbb{R}$ defined by

$$G_\beta(\boldsymbol{x}, m) := m + \frac{1}{1-\beta}\mathbb{E}\left[[\mathcal{L}(\boldsymbol{x}, \tilde{\boldsymbol{\xi}}) - m]^+ \right]. \tag{2.30}$$

They show that G_β is convex with respect to m and provide the following characterization of minimizing $\mathrm{CVaR}_\beta(\boldsymbol{x})$ in terms of the auxiliary function:

$$\underset{\boldsymbol{x}\in X}{\text{minimize}}\ \mathrm{CVaR}_\beta(\boldsymbol{x}) = \underset{(\boldsymbol{x},m)\in X\times\mathbb{R}}{\text{minimize}}\ G_\beta(\boldsymbol{x}, m), \tag{2.31}$$

This relation shows that the minimal value $\mathrm{CVaR}_\beta(\boldsymbol{x}^*)$ can be achieved by minimizing the function $G_\beta(\boldsymbol{x}, m)$ with respect to $\boldsymbol{x} \in X$ and $m \in \mathbb{R}$, simultaneously. Furthermore, it is shown in [74] that, with an optimal solution $(\boldsymbol{x}^*, m^*)$ of the right-hand side optimization problem in (2.31), $\boldsymbol{x}^*$ is an optimal solution of the left-hand side one, and m^* is close (sometimes equal) to $\mathrm{VaR}_\beta(\boldsymbol{x}^*)$.

Proposition 2.5 (Rockafellar and Uryasev [74]) *The function (2.30) is convex in $(\boldsymbol{x}, m)$ if the loss function $\mathcal{L}(\cdot, \boldsymbol{\xi})$ from $\mathbb{R}^J$ to $(-\infty, \infty]$ is convex.* ⋆

We assume that there is a finite number of realizations of random variable $\tilde{\boldsymbol{\xi}}$ and denote by $\boldsymbol{\xi}_s$ a realization and by p_s the occurrence probability of $s \in \mathcal{S}$. Then, β-CVaR is evaluated by the optimal value of the following optimization problem

[74]:

$$\left|\begin{array}{ll} \underset{(\boldsymbol{x},m,\boldsymbol{\tau})\in X\times \mathbb{R}\times \mathbb{R}^S}{\text{minimize}} & m + \dfrac{1}{1-\beta}\displaystyle\sum_{s\in S} p_s\tau_s \\ \text{subject to} & \tau_s \geq \mathcal{L}(\boldsymbol{x},\boldsymbol{\xi}_s) - m, \quad \tau_s \geq 0, \quad s \in \mathcal{S}. \end{array}\right. \tag{2.32}$$

When the loss function $\mathcal{L}(\cdot,\boldsymbol{\xi})$ is linear, Problem (2.32) is a linear programming problem and can be solved efficiently.

In the following, we will explain stochastic dominance and coherent risk measure, showing that CVaR has some theoretically desirable properties of risk measure which VaR doesn't exhibit (e.g., [69, 74]).

Stochastic Dominance *Stochastic dominance* refers to the relations that may hold between a pair of random variables. A random variable $\tilde{X}$ is said to dominate another random variable $\tilde{Y}$ in the first order, which we denote by $\tilde{X} \succeq_{(1)} \tilde{Y}$, if

$$F_{\tilde{X}}(\eta) \leq F_{\tilde{Y}}(\eta) \quad \text{for all } \eta \in \mathbb{R}, \tag{2.33}$$

where $F_{\tilde{X}}(\eta) = \text{Prob}\{\tilde{X} \leq \eta\}$ denotes the distribution function of a random variable $\tilde{X}$. It is known that the first order stochastic dominance relation $\tilde{X} \succeq_{(1)} \tilde{Y}$ is equivalent to the relation

$$\mathbb{E}[u(\tilde{X})] \geq \mathbb{E}[u(\tilde{Y})] \quad \text{for all non-decreasing functions } u : \mathbb{R} \to \mathbb{R}. \tag{2.34}$$

The second order stochastic dominance relation $\tilde{X} \succeq_{(2)} \tilde{Y}$ is defined as

$$\int_{-\infty}^{\eta} F_{\tilde{X}}(\phi)\,\mathrm{d}\phi \leq \int_{-\infty}^{\eta} F_{\tilde{Y}}(\phi)\,\mathrm{d}\phi \quad \text{for all } \eta \in \mathbb{R}, \tag{2.35}$$

or equivalently,

$$\mathbb{E}[u(\tilde{X})] \geq \mathbb{E}[u(\tilde{Y})] \quad \text{for all non-decreasing concave functions } u : \mathbb{R} \to \mathbb{R}. \tag{2.36}$$

The risk-averse utility function is non-decreasing and concave, therefore risk-averse expected-utility maximizers prefer a random return $\tilde{X}$ to $\tilde{Y}$ when $\tilde{X} \succeq_{(2)} \tilde{Y}$.

Proposition 2.6 (Pflug [69]) *Suppose that $\tilde{X} \succeq_{(1)} \tilde{Y}$ for two random variables $\tilde{X}$ and $\tilde{Y}$. Then the β-VaR of $\tilde{X}$ is no less than the β-VaR of $\tilde{Y}$, and the β-CVaR of $\tilde{X}$ is no less than the β-CVaR of $\tilde{Y}$.*

For two random variables $\tilde{X}$ and $\tilde{Y}$ if $\tilde{X} \succeq_{(2)} \tilde{Y}$ then the β-CVaR of $\tilde{X}$ is no less than the β-CVaR of $\tilde{Y}$. ⋆

Coherent Risk Measure We consider a random loss $\tilde{Z} = Z(\omega)$ being an element of a linear space $\mathcal{Z}$ of $\mathcal{F}$-measurable functions defined on the sample space $(\Omega, \mathcal{F})$. A risk measure ρ is said to be *coherent risk measure* for $\mathcal{Z}$ if it satisfies the following conditions:

Axiom 2.7 (Convexity)
$\rho(\lambda\tilde{Z}_1 + (1-\lambda)\tilde{Z}_2) \leq \lambda\rho(\tilde{Z}_1) + (1-\lambda)\rho(\tilde{Z}_2)$ for all $\tilde{Z}_1, \tilde{Z}_2 \in \mathcal{Z}$ and all $\lambda \in [0,1]$. ⋆

Axiom 2.8 (Monotonicity)
If $\tilde{Z}_1, \tilde{Z}_2 \in \mathcal{Z}$ and $Z_1(\omega) \geq Z_2(\omega)$ for a.e. $\omega \in \Omega$, then $\rho(\tilde{Z}_1) \geq \rho(\tilde{Z}_2)$. ⋆

Axiom 2.9 (Translation Equivalence)
If $\alpha \in \mathbb{R}$ and $\tilde{Z} \in \mathcal{Z}$, then $\rho(\tilde{Z} + \alpha) = \rho(\tilde{Z}) + \alpha$. ⋆

Axiom 2.10 (Positive Homogeneity)
If $\lambda > 0$ and $\tilde{Z} \in \mathcal{Z}$, then $\rho(\lambda\tilde{Z}) = \lambda\rho(\tilde{Z})$. ⋆

These conditions describe the concept of rational investors under uncertainty, and therefore desirable risk measures should satisfy them.

Proposition 2.11 (Pflug [69]) *β-CVaR is a coherent risk measure.* ⋆

In Chapter 3, we apply the CVaR to the classical newsvendor problem and devise solution method.

2.4.2 Portfolio Optimization

Portfolio optimization problem is to determine an investment proportion of initial wealth into financial assets. In general, the return of financial assets includes uncertainty, and accordingly risk measures are considered for averting a large loss of portfolio value.

Markowitz [53] is best known for his pioneering work in Modern Portfolio Theory. Portfolio optimization problem is usually formulated via two-parameter approach as in the Markowitz's mean-variance model [53] where the two parameters represent expected value and variance of a portfolio return:

$$
\left|
\begin{array}{ll}
\underset{\boldsymbol{w}\in\mathbb{R}^J}{\text{minimize}} & \boldsymbol{w}^\top \boldsymbol{\Sigma} \boldsymbol{w} \\
\text{subject to} & \bar{\boldsymbol{r}}^\top \boldsymbol{w} = R \\
& \boldsymbol{w} \in W,
\end{array}
\right.
\tag{2.37}
$$

where

$\boldsymbol{w}$	:	investment proportion (decision variables), $\boldsymbol{w} \in \mathbb{R}^J$
$\boldsymbol{\Sigma}$	:	covariance matrix of the random return $\tilde{\boldsymbol{r}}$ of financial assets (given), $\boldsymbol{\Sigma} \in \mathbb{R}^{J\times J}$
$\bar{\boldsymbol{r}}$	:	expected return of financial assets (given), $\bar{\boldsymbol{r}} \in \mathbb{R}^J$
R	:	target return (given), $R \in \mathbb{R}$
W	:	set of feasible investment proportions, $W \subseteq \{\boldsymbol{w} \in \mathbb{R}^J \mid \boldsymbol{1}^\top \boldsymbol{w} = 1\}$

Although Problem (2.37) is convex quadratic programming problem, it becomes intractable when the number of investable assets, J grows. Suppose that we have a sample of S realizations of the random vector $\tilde{\boldsymbol{r}}$, and denote by $\boldsymbol{r}_s$ the realizations and by p_s the occurrence probability of $s \in \mathcal{S}$. Then compact factorization of the mean-variance model by Konno and Suzuki [44] is as follows:

$$
\left|
\begin{array}{ll}
\underset{(\boldsymbol{w},\boldsymbol{z})\in\mathbb{R}^J\times\mathbb{R}^S}{\text{minimize}} & \sum_{s\in\mathcal{S}} p_s z_s^2 \\
\text{subject to} & z_s - (\boldsymbol{r}_s - \bar{\boldsymbol{r}})^\top \boldsymbol{w} = 0, \quad s \in \mathcal{S} \\
& \bar{\boldsymbol{r}}^\top \boldsymbol{w} = R \\
& \boldsymbol{w} \in W.
\end{array}
\right.
\tag{2.38}
$$

Problem (2.38) has quadratic terms as many as S. Therefore it is solved efficiently if the number of scenarios S is not very large.

Konno and Yamazaki [43] also proposed a mean-absolute deviation model:

$$
\left|
\begin{array}{ll}
\underset{(\boldsymbol{w},\boldsymbol{y})\in\mathbb{R}^J\times\mathbb{R}^S}{\text{minimize}} & \sum_{s\in\mathcal{S}} p_s |y_s| \\
\text{subject to} & y_s - (\boldsymbol{r}_s - \bar{\boldsymbol{r}})^\top \boldsymbol{w} = 0, \quad s \in \mathcal{S} \\
& \bar{\boldsymbol{r}}^\top \boldsymbol{w} = R \\
& \boldsymbol{w} \in W,
\end{array}
\right.
\tag{2.39}
$$

where the quadratic terms are replaced by absolute values of y_s. Although Problem (2.39) is a non-differentiable optimization problem which seems to be more difficult than the mean-variance model, Problem (2.39) is reformulated as the following linear programming problem:

$$\left|\begin{array}{ll} \underset{(\boldsymbol{w},\boldsymbol{y})\in\mathbb{R}^J\times\mathbb{R}^S}{\text{minimize}} & \displaystyle\sum_{s\in\mathcal{S}} p_s y_s \\ \text{subject to} & y_s - (\boldsymbol{r}_s - \bar{\boldsymbol{r}})^\top \boldsymbol{w} \geq 0, \quad s\in\mathcal{S} \\ & y_s + (\boldsymbol{r}_s - \bar{\boldsymbol{r}})^\top \boldsymbol{w} \geq 0, \quad s\in\mathcal{S} \\ & \bar{\boldsymbol{r}}^\top \boldsymbol{w} = R \\ & \boldsymbol{w} \in W. \end{array}\right. \tag{2.40}$$

Among various risk measures examined in the literature (see, e.g., [22]), we employ the VaR and the CVaR in this thesis. The random loss regarding the VaR and CVaR is defined as $-\bar{\boldsymbol{r}}^\top \boldsymbol{w}$. Mean-VaR model is formulated as follows:

$$\left|\begin{array}{ll} \underset{(\boldsymbol{w},m)\in\mathbb{R}^J\times\mathbb{R}}{\text{minimize}} & m \\ \text{subject to} & \text{Prob}\{-\tilde{\boldsymbol{r}}^\top \boldsymbol{w} \leq m\} \geq \beta \\ & \bar{\boldsymbol{r}}^\top \boldsymbol{w} = R \\ & \boldsymbol{w} \in W, \end{array}\right. \tag{2.41}$$

and mean-CVaR model is formulated as the following linear programming problem:

$$\left|\begin{array}{ll} \underset{(\boldsymbol{w},m,\boldsymbol{\tau})\in X\times\mathbb{R}\times\mathbb{R}^S}{\text{minimize}} & m + \dfrac{1}{1-\beta}\displaystyle\sum_{s\in S} p_s \tau_s \\ \text{subject to} & \tau_s \geq -\boldsymbol{r}_s^\top \boldsymbol{w} - m, \quad \tau_s \geq 0, \quad s\in\mathcal{S}. \end{array}\right. \tag{2.42}$$

where m is a decision variable corresponding to the Value-at-Risk. Problem (2.41) is an optimization problem with a probabilistic constraint (see Section 2.2), which means this problem is nonconvex and intractable in general.

In Chapter 4, we propose a solution method for VaR minimization problem of a financial asset portfolio by applying a branch-and-bound procedure, and in Chapter 5, we solve a multi-period portfolio optimization problem where the CVaR constraints are taken into account for avoiding a large loss of portfolio value.

Chapter 3

Newsvendor Solutions via Conditional Value-at-Risk Minimization

3.1 Newsvendor Problem in Single Period

The classic newsvendor problem is to decide the quantity of the product whose demand is uncertain by maximizing the expected profit, or equivalently minimizing the expected cost. In this section, we briefly summarize the classic single-period newsvendor problem for the comparison with our results.

Notation

First of all, let us introduce notation used in this chapter as follows:

$\mathcal{N}$: index set for products, $\mathcal{N} := \{1, 2, ..., N\}$,

$\tilde{\xi}_i$: daily demand for product i (non-negative, scalar-valued random variable)

q_i : selling price per unit for product i (given)

c_i : cost per unit for product i (given)

r_i : salvage value per unit for product i (given)

π_i : shortage penalty per unit for product i (given)

x_i : daily order quantity for product i (decision variable).

We assume the following condition:

Assumption 3.1 $r_i < c_i < q_i, \;\; \pi_i \geq 0$ *for all* $i \in \mathcal{N}$. $\star$

In the following, we omit the subscript for simplicity when only single product is considered.

Profit Maximization and Cost Minimization

For fixed x, the daily profit gained from each product is a random variable defined by

$$\mathcal{P}(x, \tilde{\xi}) := q \min \left\{ \tilde{\xi},\, x \right\} + r \max \left\{ x - \tilde{\xi},\, 0 \right\} - \pi \max \left\{ \tilde{\xi} - x,\, 0 \right\} - c\, x, \tag{3.1}$$

where the third term in the right-hand side represents an artificial penalty for opportunity cost, and π is often set to be 0.

Let $F(\eta)$ denote the distribution function of demand for the product, i.e., $F(\eta) := \text{Prob}\left\{ \tilde{\xi} \leq \eta \right\}$. We note that $F(0) = 0$. The classic newsvendor model then maximizes the expected profit:

$$\left| \; \underset{x}{\text{maximize}} \;\; \mu(x) := \mathbb{E}\left[\mathcal{P}(x, \tilde{\xi}) \right] = \int_0^\infty \mathcal{P}(x, \xi)\, \mathrm{d}F(\xi). \right. \tag{3.2}$$

When the inverse of the distribution function exists, an optimal solution of Problem (3.2) is obtained by solving $\frac{\partial \mu}{\partial x} = 0$, as

$$x^* = F^{-1}\!\left(\frac{U}{E + U} \right), \tag{3.3}$$

where $E := c - r$, and $U := q + \pi - c$. Even when F does not have the inverse, one can obtain a solution via a simple numerical calculation [77].

We also formulate Problem (3.2) as a two-stage problem as follows:

$$\left| \; \underset{x \in \mathbb{R}}{\text{minimize}} \quad cx + \mathbb{E}[\tilde{Q}(x)], \right. \tag{3.4}$$

where $\tilde{Q}(x)$ is defined by an optimal value of the following optimization problem

$$\left| \; \begin{array}{ll} \underset{\boldsymbol{y} \in \mathbb{R}^3}{\text{minimize}} & -q\, y_1 - r\, y_2 + \pi\, y_3 \\ \text{subject to} & y_1 \leq \tilde{\xi} \\ & y_1 + y_2 \leq x \\ & y_3 \geq \tilde{\xi} - x \\ & \boldsymbol{y} \geq \mathbf{0}. \end{array} \right. \tag{3.5}$$

On the other hand, the daily total cost is defined by

$$\mathcal{Q}(x, \tilde{\xi}) := E\,[x - \tilde{\xi}]^+ + U\,[\tilde{\xi} - x]^+. \tag{3.6}$$

Here, the first term in the right-hand side of (3.6) represents the cost for excess order, while the second does the opportunity cost. By noting the relation

$$\mathcal{P}(x, \tilde{\xi}) = V\,\tilde{\xi} - \mathcal{Q}(x, \tilde{\xi}), \tag{3.7}$$

where $V := q - c = U - \pi$, the minimization of the expected cost is proved to be equivalent to the maximization of the expected profit:

$$\min_x \mathbb{E}\left[\,\mathcal{Q}(x, \tilde{\xi})\,\right] \;=\; V\,\mathbb{E}\left[\,\tilde{\xi}\,\right] \;-\; \max_x \mathbb{E}\left[\,\mathcal{P}(x, \tilde{\xi})\,\right].$$

Since $E, U > 0$ from Assumption 3.1, the expected cost is a convex function in x, whereas the expected profit is concave one, and therefore, both problems are so-called convex programs.

In the case where multiple products are considered, we assume that the total profit (or cost) is just the sum of the ones from each product, i.e., the total profit $\mathcal{P}(\boldsymbol{x}, \tilde{\boldsymbol{\xi}})$ and the total cost $\mathcal{Q}(\boldsymbol{x}, \tilde{\boldsymbol{\xi}})$ are given by

Assumption 3.2 $\mathcal{P}(\boldsymbol{x}, \tilde{\boldsymbol{\xi}}) = \sum_{i \in \mathcal{N}} \mathcal{P}(x_i, \tilde{\xi}_i);\quad \mathcal{Q}(\boldsymbol{x}, \tilde{\boldsymbol{\xi}}) = \sum_{i \in \mathcal{N}} \mathcal{Q}(x_i, \tilde{\xi}_i).$ $\star$

3.2 Minimization of CVaR in the Newsvendor Problem

In this section, we show that the CVaR minimization in the newsvendor situation leads to a convex problem when the associated loss is a convex function. We will apply two different functions as the loss and derive closed form solutions in the succeeding sections. Also, the parameter sensitivity of the solution is examined.

In the following, we consider two different loss functions. One is defined by $-\mathcal{P}(\boldsymbol{x}, \tilde{\boldsymbol{\xi}})$ and called the *net loss* of the profit, while the other one is the *total cost* $\mathcal{Q}(\boldsymbol{x}, \tilde{\boldsymbol{\xi}})$. Since both the net loss, $-\mathcal{P}(\cdot, \boldsymbol{\xi})$, and the total cost, $\mathcal{Q}(\cdot, \boldsymbol{\xi})$, are convex functions for fixed $\boldsymbol{\xi}$ under Assumptions 3.1 and 3.2, the CVaR minimization problems using these functions are convex from Proposition 2.5. In the following,

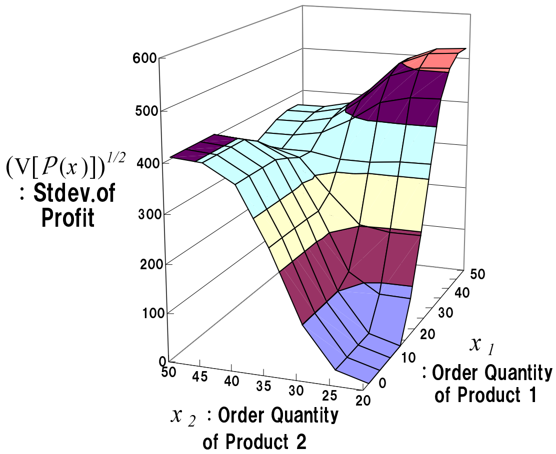


Figure 3.1: Nonconvexity of Standard Deviation of Profit $\mathcal{P}$

we call the CVaR minimization problems using these functions the net loss CVaR minimization and the total cost CVaR minimization, respectively.

It is worth noting that the convexity-preserving property is also valid for a class of downside risk measures.

Proposition 3.3 (Theorem 5.1, Rockafellar [73]) *Let g be a convex function from $\mathbb{R}^n$ to $(-\infty, \infty]$, and let γ be a convex function from $(-\infty, \infty]$ to $(-\infty, \infty]$ which is non-decreasing with $\gamma(\infty) = \infty$. Then, $h(x) = \gamma(g(x))$ is convex on $\mathbb{R}^n$.* ⋆

Proof : See [73], for example. ■

From this proposition, we see that minimization of any non-decreasing convex risk measure including the below-target return defined by $\mathbb{E}[[t-\mathcal{P}(\boldsymbol{x}, \tilde{\boldsymbol{\xi}})]^+]$ for fixed target $t \in \mathbb{R}$ [31] and the maximal loss, $\max_{\boldsymbol{\xi}}\{-\mathcal{R}(\boldsymbol{x}, \boldsymbol{\xi}) \mid \boldsymbol{\xi} \in \Xi\}$, which is defined when $\tilde{\boldsymbol{\xi}}$ has finite support Ξ [83], is formulated as a convex problem, and at the same time, the lower partiality of the risk measures seems crucial for the convexity in the risk minimization for the newsvendor problem. In fact, the variance (or equivalently, the standard deviation) of the net loss or the total cost function can have a non-convex structure. Figure 3.1 shows an example of the non-convexity with respect to $\boldsymbol{x}$ in the standard deviation of profit $\mathcal{P}$ in the two-product case where the underlying distribution has finite supports.

3.2.1 Unconstrained Minimization of CVaR for Single Product Case

In this subsection, we present analytical results of the CVaR minimization problems for the case of single product without constraint. Let us assume for simplicity that there exists the inverse F^{-1} of the distribution function F of demand $\tilde{\xi}$, and denote its density by f.

a) The Net Loss CVaR Minimization First, we adopt the net loss $-\mathcal{P}$ as the loss function $\mathcal{L}$ so that a manager can consider the profit lower than $-\mathrm{VaR}_\beta$.

The minimization of (2.30) with $\mathcal{L} = -\mathcal{P}$ is represented as the following convex program:

$$\left| \underset{x \in \mathbb{R}, m \in \mathbb{R}}{\text{minimize}} \quad p(x,m) := m + \frac{1}{1-\beta} \int_0^\infty [-\mathcal{P}(x,\xi) - m]^+ \, f(\xi)\, \mathrm{d}\xi. \right. \tag{3.8}$$

This problem can be solved in a closed form under mild assumption as shown in the following proposition, whose proof is given in Appendix A.

Proposition 3.4 *The problem (3.8) with $\beta \in [0,1)$ has an optimal solution (x^*, m^*) defined by*

$$\begin{cases} x^* = \dfrac{E+V}{E+U} F^{-1}\!\left(\dfrac{U(1-\beta)}{E+U} \right) + \dfrac{U-V}{E+U} F^{-1}\!\left(\dfrac{E\beta + U}{E+U} \right), \\ m^* = \dfrac{E\,(U-V)}{E+U} F^{-1}\!\left(\dfrac{E\beta+U}{E+U} \right) - \dfrac{U\,(E+V)}{E+U} F^{-1}\!\left(\dfrac{U(1-\beta)}{E+U} \right). \end{cases} \tag{3.9}$$

When $\beta = 0$, any m^* with $m^* \leq -Vx^*$ satisfies the optimality. $\star$

In particular, when the shortage penalty π is set to be 0, i.e., $V = U$, we have the following simpler result:

Corollary 3.5 *Under the same assumption as in Proposition 3.4 with $\pi = 0$, we have an optimal solution (x^*, m^*) defined by*

$$x^* = F^{-1}\left(\frac{U}{E+U} (1-\beta) \right); \quad m^* = -Ux^*. \tag{3.10}$$

When $\beta = 0$, any m^* with $m^* \leq -Ux^*$ satisfies the optimality.

From these proposition and corollary, we see that the difference between the solution x^* given by (3.9) or (3.10) and the classic one (3.3) depends on two parameters π and β. In particular, we see from Corollary 3.5 that when $\pi = 0$, they differ only in the coefficient of the argument of F^{-1}, whereas when $\pi > 0$, it is much more complex. This CVaR minimization gives a simple generalization of the classic problem since the solutions with $\beta = 0$ is equal to the classic one (3.3). Moreover, the availability of the closed form solution plays a role in sensitivity analysis.

b) The Total Cost CVaR Minimization Next, we consider the total cost $\mathcal{Q}$ as the loss $\mathcal{L}$. By minimizing the β-CVaR defined for the total cost, a manager may avoid an unduly large cost which made of the excess order cost and the opportunity cost.

The corresponding problem is

$$\left| \begin{array}{l} \underset{x \in \mathbb{R}, m \in \mathbb{R}}{\text{minimize}} \quad q(x,m) := m + \frac{1}{1-\beta} \int_0^\infty [\, \mathcal{Q}(x,\xi) - m\,]^+ \, f(\xi)\, \mathrm{d}\xi. \end{array} \right. \tag{3.11}$$

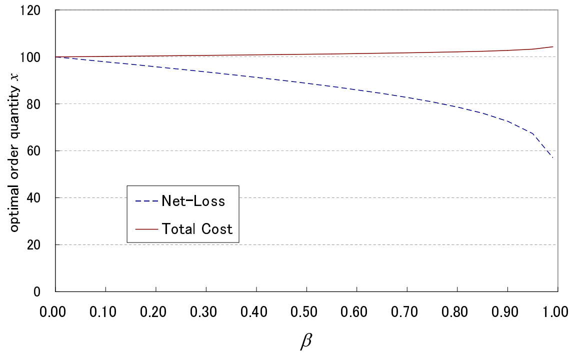
Since the total cost $\mathcal{Q}$ can be treated mathematically as a special case of the net loss $-\mathcal{P}$ with $V = 0$, we obtain a solution of (3.11) from Proposition 3.4 by setting $V = 0$.

Corollary 3.6 *The problem (3.11) with $\beta \in [0,1)$ has an optimal solution (x^*, m^*) defined by*

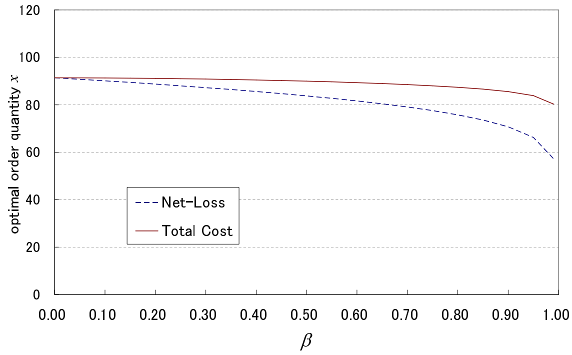
$$\left\{ \begin{array}{l} x^* = \dfrac{E}{E+U} F^{-1}\left(\dfrac{U(1-\beta)}{E+U} \right) + \dfrac{U}{E+U} F^{-1}\left(\dfrac{E\beta + U}{E+U} \right), \\ m^* = \dfrac{EU}{E+U} \left(F^{-1}\left(\dfrac{E\beta+U}{E+U} \right) - F^{-1}\left(\dfrac{U(1-\beta)}{E+U} \right) \right). \end{array} \right. \tag{3.12}$$

When $\beta = 0$, any m^* with $m^* \leq 0$ satisfies the optimality.

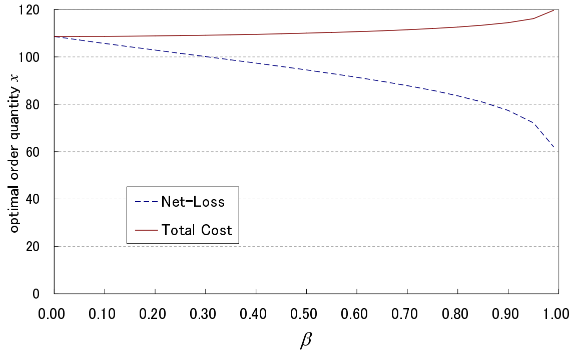
By comparing solutions (3.9) and (3.12) of the two CVaR minimizations, we observe that the solution of the total cost CVaR minimization (3.11) can be far different from that of the net loss CVaR minimization (3.8), whereas maximizing the profit and minimizing the cost are equivalent in the classic problem (see the results in Section 3.1). However, the CVaR minimization (3.11) also provides a generalization of the classic maximizing profit model because the solution (3.12) with $\beta = 0$ is the same as the solution (3.3).



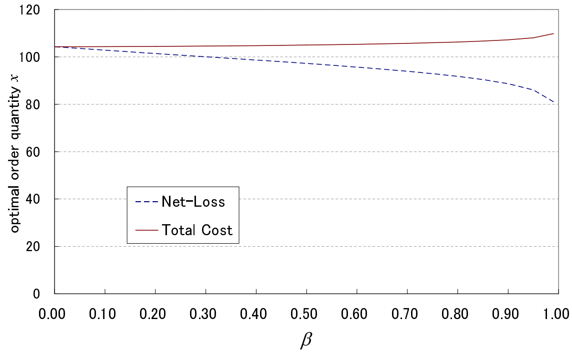
(a) $(E, U, V) = (60, 60, 50)$; $\mathrm{N}(100, 20^2)$



(b) $(E, U, V) = (60, 30, 20)$; $\mathrm{N}(100, 20^2)$



(c) $(E, U, V) = (30, 60, 50)$; $\mathrm{N}(100, 20^2)$



(d) $(E, U, V) = (30, 60, 50)$; $\mathrm{N}(100, 10^2)$

Figure 3.2: Optimal Solutions of Two CVaR Minimizations with Different Loss Functions

Figures 3.2 (a) to (d) illustrate the differences among the three optima (3.3), (3.9) and (3.12) when π is set to be 10, i.e., $V = U - 10$, and ξ follows a normal distribution with different parameter settings. Though the normal distribution is not adequate for describing the product demand since $F(0) > 0$, we here apply it in order to roughly grasp the dependency of the solutions on the shape of distribution. Noting that the solution with $\beta = 0$ is equal to the classic expected profit maximizer (3.3), we see from the figures that the net loss CVaR minimization implies smaller order quantity than the classic solution and the difference becomes larger as β gets higher. The optimal solution of the total cost CVaR depends on parameters E and U. In particular, when $E < U$ holds as in (a), (c) and (d), the two CVaR minimizers with different loss functions show reverse trends with β. Also, from Figures 3.2 (c) and (d), we see that the difference between the two solutions becomes smaller as the variance of normal distribution decreases. Especially, the paces of decrease of the order quantity based on the net loss CVaR minimization and increase of that on the total cost CVaR minimization are more than linear with respect to β. This suggests that the value of β and the loss function should be cautiously chosen.

3.2.2 Sensitivity Analysis

As illustrated in Figure 3.2, the solutions (3.9) and (3.12) depend on the parameters β and π. In order to clarify how the CVaR minimizers and the classic expected profit maximizer (3.3) disagree, we here analyze the parameter sensitivity of the solutions.

The signs of $\frac{\partial x^*}{\partial \beta}$ for the net loss and the total cost CVaR minimizers depend on the underlying distribution because we have for the net loss CVaR minimizer (3.9) with $\pi = U - V > 0$,

$$\frac{\partial x^*}{\partial \beta} = \frac{1}{(E+U)^2}\left(\frac{E(U-V)}{f(F^{-1}(G_2))} - \frac{U(E+V)}{f(F^{-1}(G_1))}\right), \tag{3.13}$$

where $G_1 = \frac{U(1-\beta)}{E+U}$ and $G_2 = \frac{E\beta+U}{E+U}$, and for the total cost CVaR minimizer (3.12),

$$\frac{\partial x^*}{\partial \beta} = \frac{EU}{(E+U)^2}\left(\frac{1}{f(F^{-1}(G_2))} - \frac{1}{f(F^{-1}(G_1))}\right).$$

The sign of $\frac{\partial x^*}{\partial q}$ for the the net loss CVaR minimizer depends also on the distribution when $\pi > 0$ since

$$\frac{\partial x^*}{\partial q} = \frac{U-V}{(E+U)^2}\left(F^{-1}(G_1) - F^{-1}(G_2)\right) + \frac{(1-\beta)E}{(E+U)^3}\left(\frac{E+V}{f(F^{-1}(G_1))} + \frac{U-V}{f(F^{-1}(G_2))}\right).$$

To illustrate how the shape of distribution affects the derivatives, let us consider the S-D distribution which has been used to analyze the newsvendor problem because of its tractability (e.g., [46, 50]). The distribution function and density function of S-D distribution are defined, respectively, by

$$F(\eta) = \begin{cases} d - \{(a-\eta)/b\}^{\frac{1}{l}}, & \text{for } \eta \in [H_1, a), \\ d + \{(\eta - a)/b\}^{\frac{1}{l}}, & \text{for } \eta \in [a, H_2], \end{cases}$$

and

$$f(\eta) = \frac{1}{bl}\left|\frac{a-\eta}{b}\right|^{\frac{1-l}{l}}, \quad \text{for } H_1 \le \eta \le H_2,$$

Table 3.1: Sign of Partial Derivative of Each Minimizer

	$\frac{\partial x^*}{\partial q}$	$\frac{\partial x^*}{\partial c}$	$\frac{\partial x^*}{\partial r}$	$\frac{\partial x^*}{\partial \pi}$	$\frac{\partial x^*}{\partial \beta}$
net loss CVaR ($\pi = 0$) (3.10)	$+$	$-$	$+$	0	$-$
net loss CVaR ($\pi > 0$) (3.9)	case-by-case	$-$	$+$	$+$	case-by-case
total cost CVaR (3.12)	$+$	$-$	$+$	$+$	case-by-case
Classical (3.3)	$+$	$-$	$+$	$+$	0

where a, b, d, l, H_1 and H_2 are constant with $b > 0$, $l > 0$ and $d \in \left[\{(a - H_1)/b\}^{1/l}, 1 - \{(H_2 - a)/b\}^{1/l}\right]$. To focus on the difference between the two CVaR minimizers, we consider the case of $l \in (0, 1)$. For the net loss CVaR minimizer, we have

$$\begin{cases} \dfrac{\partial x^*}{\partial \beta} \geq 0 \quad \text{if} \quad d \in [(1-\theta)G_1 + \theta G_2, (1-\nu)G_1 + \nu G_2), \\ \dfrac{\partial x^*}{\partial \beta} \leq 0 \quad \text{if} \quad d < (1-\theta)G_1 + \theta G_2 \quad \text{or} \quad d \geq (1-\nu)G_1 + \nu G_2, \end{cases}$$

where $\nu := B^{1/(1-l)}/(B^{1/(1-l)} - 1)$, $\theta := B^{1/(1-l)}/(B^{1/(1-l)} + 1)$ and $B := \frac{U(E+V)}{E(U-V)}$. On the other hand, for the total cost CVaR solution, we then have

$$\frac{\partial x^*}{\partial \beta} \geq 0 \quad \text{if} \quad d \geq \frac{G_1 + G_2}{2}; \qquad \frac{\partial x^*}{\partial \beta} < 0 \quad \text{if} \quad d < \frac{G_1 + G_2}{2}.$$

From the above results, we see that the sensitivity with respect to β depends on the parameters of the S-D distribution. In particular, the skewness parameter d is crucial for the solutions above. Moreover, we see that these two derivatives can show different signs for the same parameter setting.

Table 3.1 summarizes the sign of the partial derivative of each optimal solution x^* with respect to parameters q, c, r, π and β. All the signs of the sensitivity to c, r and π remain the same as that of the classic expected profit maximization model, whereas those to β and q can differ from model to model. In particular, when $\pi \geq 0$ is sufficiently small, we see from the table and Equation (3.13) that the net loss CVaR minimizer provides smaller order quantity than the classic expected profit maximizer (3.3) for $\beta > 0$. This implies that a risk-averse manager who cares about large loss via the β-CVaR is likely to order smaller quantity than the risk neutral solution (3.3).

3.3 Mean-CVaR Models and LP Formulation

Since the Markowitz's seminal work, the trade-off between risk and return has been considered and optimized in various situations. This trade-off model is known as the *mean-risk model* (see [64], for example), which is formulated as the optimization of a composite objective consisting of the expected return and a certain risk measure ρ:

$$\left|\begin{array}{ll} \underset{\boldsymbol{x}}{\text{maximize}} & \mathbb{E}\left[\mathcal{P}(\boldsymbol{x},\tilde{\boldsymbol{\xi}})\right] - \lambda\,\rho(\mathcal{L}(\boldsymbol{x},\tilde{\boldsymbol{\xi}})) \\ \text{subject to} & \boldsymbol{x} \in X, \end{array}\right. \tag{3.14}$$

where X a convex set representing some constraints on the portfolio $\boldsymbol{x}$, and $\lambda \geq 0$ a trade-off parameter, or formulated as the minimization of the risk, keeping the return at least as large as a predetermined target:

$$\left|\begin{array}{ll} \underset{\boldsymbol{x}}{\text{minimize}} & \rho(\mathcal{L}(\boldsymbol{x},\tilde{\boldsymbol{\xi}})) \\ \text{subject to} & \mathbb{E}\left[\mathcal{P}(\boldsymbol{x},\tilde{\boldsymbol{\xi}})\right] \geq \overline{\mu}, \\ & \boldsymbol{x} \in X, \end{array}\right. \tag{3.15}$$

where $\overline{\mu}$ is the minimum level of the expected profit. It is known that the both formulations give the same convex efficient frontier, which is a graph of Pareto efficient pairs of expected return and some risk measure ρ, when the expected return is a concave function of $\boldsymbol{x}$ and the risk is a convex one. Exploiting the results in the previous section and applying the CVaR measures as the risk ρ, the corresponding mean-risk models (3.14) and (3.15) are convex programs, and result in the same efficient frontier.

3.3.1 Unconstrained Mean-CVaR Models for Single Product Case

The unconstrained mean-risk model using the net loss CVaR is formulated as

$$\left|\ \underset{x,m}{\text{maximize}} \quad \int_0^\infty \mathcal{P}(x,\xi)\,f(\xi)\,\mathrm{d}\xi - \lambda\left(m + \frac{1}{1-\beta}\int_0^\infty \left[\,-\mathcal{P}(x,\xi) - m\,\right]^+ f(\xi)\,\mathrm{d}\xi\right).\right. \tag{3.16}$$

By the same reasoning as in the net loss CVaR minimization developed in Section 3.2.1, we consider the following three cases in order to evaluate the integral part of the CVaR measure (see the proof of Proposition 3.4 in Appendix A for details). Throughout the following analysis (except for proposition), we omit the case of $\lambda = 0$, since the mean-risk model is then equal to the profit maximization (3.2).

$\langle\langle$ **case 1.** $m \leq -Vx$ $\rangle\rangle$ Let us denote the objective function of (3.16) by $h(x) := \mu(x) - \lambda\, p(x, m)$ from (3.2) and (3.8). When $\beta = 0$ holds, we have the following first-order condition:

$$\begin{cases} \dfrac{\partial h}{\partial x} = -\left(1 + \dfrac{\lambda}{1-\beta}\right)\{(\,E + U\,)\,F(x) - U\} = 0, \\ \dfrac{\partial h}{\partial m} = -\lambda\left(1 - \dfrac{1}{1-\beta}\right) = 0; \end{cases} \tag{3.17}$$

we then achieve a solution (x^*, m^*) satisfying (3.3) and $m^* \leq -Vx^*$.

$\langle\langle$ **case 2.** $m \in (-Vx, Ex)$ $\rangle\rangle$ When $\pi > 0$, we have the first-order condition as

$$\begin{cases} E\left\{F(x) + \dfrac{\lambda}{1-\beta}F\left(\dfrac{Ex - m}{E + V}\right)\right\} \\ \qquad + U\left\{F(x) + \dfrac{\lambda}{1-\beta}F\left(\dfrac{Ux + m}{U - V}\right)\right\} = U\left(1 + \dfrac{\lambda}{1-\beta}\right), \\ F\left(\dfrac{Ux + m}{U - V}\right) = F\left(\dfrac{Ex - m}{E + V}\right) + \beta. \end{cases}$$

The term $F\left(\frac{Ux+m}{U-V}\right)$ can be eliminated from the first equation by substituting the second one. Thus, an optimal m should satisfy

$$m = Ex - (\,E + V\,)F^{-1}(\,A(x)\,), \tag{3.18}$$

where $A(x) := \frac{1-\beta}{\lambda}\left\{\frac{U}{E+U}(\,1 + \lambda\,) - F(x)\right\}$. By substituting m defined by (3.18) into the second equation of the optimality condition, we see that optimal x should satisfy

$$(\,E + V\,)\,F^{-1}(A(x)) + (\,U - V\,)\,F^{-1}(A(x) + \beta) - (\,E + U\,)x = 0. \tag{3.19}$$

Note that the left-hand side of (3.19) is decreasing with respect to x, so it is not hard to compute an optimal order quantity satisfying the equality.

$\langle\langle$ **case 3.** $m \geq Ex$ $\rangle\rangle$ By solving the first-order condition, we have a solution (x^*, m^*) defined by

$$x^* = F^{-1}\left(\frac{U}{E + U}(1 + \lambda)\right); \quad m^* = (\,U - V\,)\,F^{-1}(\beta) - U\,x^*. \tag{3.20}$$

Considering the condition $m \geq Ex$, β and λ should satisfy the relation

$$F^{-1}\left(\frac{U}{E+U}(1+\lambda)\right) \leq \frac{U-V}{E+U}F^{-1}(\beta). \tag{3.21}$$

Therefore, for $\lambda > 0$ and $\pi > 0$, an optimal solution can be found through the following steps:

Numerical Procedure for Mean-Net Loss CVaR Model ($\lambda > 0$, $\pi > 0$)

1. If $\beta = 0$, then (x^*, m^*) satisfying (3.3) and $m^* \leq -Vx^*$ is a solution.
2. If β and λ satisfy the relation (3.21), then (x^*, m^*) satisfying (3.20) is an optimal solution.
3. Otherwise, search x satisfying (3.19), and m defined by (3.18).

In particular when $\pi = 0$ is assumed, a closed form solution of Problem (3.16) can be obtained by a similar discussion. The proof of the following proposition is given in Appendix B.

Proposition 3.7 *For $\pi = 0$, $\beta \in [0, 1)$ and $\lambda \geq 0$, the mean-CVaR model* (3.16) *has an optimal solution (x^*, m^*) defined by*

$$x^* = F^{-1}\left(\frac{U}{E+U} \cdot \frac{1+\lambda}{1+\lambda\,(1-\beta\,)^{-1}}\right); \quad m^* = -Ux^*.$$

In particular, when $\lambda = 0$ or $\beta = 0$ holds, any m^* with $m^* \leq -U\,x^*$ satisfies the optimality. Moreover, for $\lambda \in \left[0, \frac{E+U}{E}\beta - 1\right)$, a solution (x^{**}, m^{**}) defined by

$$x^{**} = F^{-1}\left(\frac{U-\lambda E}{E+U}\right); \quad m^{**} = E\,x^{**} - (\,E+U\,)\,F^{-1}(\,1-\beta\,),$$

also achieves the optimal value. In this case, so does a solution $(\hat{x}, \hat{m})$ satisfying $\hat{x} = (1-t)x^* + t\,x^{**}$ and $\hat{m} = (1-t)m^* + t\,m^{**}$ for $t \in (0, 1)$. ⋆

As for the mean-risk solution based on the total cost CVaR, we can achieve a solution scheme by setting $V = 0$ in the above three-step procedure and the associated equations referred to therein.

3.3.2 Constrained Mean-CVaR Models for Multiple Products Case

In this subsection, we address how to compute an optimal solution when multiple products are considered under a system of linear inequalities.

Suppose that the demand distribution is given by a finite number of scenarios. Let $\mathcal{S} = \{1, ..., S\}$ denote a finite index set of scenarios, and let $\mathrm{Prob}\{\tilde{\boldsymbol{\xi}} = \boldsymbol{\xi}_s\} = p_s$ for $s \in \mathcal{S}$ where $\boldsymbol{\xi}_s := (\xi_{s,1}, \xi_{s,2}, ..., \xi_{s,n})^\top$. Moreover, X is supposed to be a polytope given by $X := \{\, \boldsymbol{x} \mid \boldsymbol{C}\boldsymbol{x} \leq \boldsymbol{b} \,\}$ where $\boldsymbol{C} \in \mathbb{R}^{m \times n}$ and $\boldsymbol{b} \in \mathbb{R}^m$. Then, using the net loss CVaR, the mean-risk model (3.14) with $\rho(\mathcal{L}(\boldsymbol{x}, \tilde{\boldsymbol{\xi}})) = \mathrm{CVaR}_\beta(\boldsymbol{x})$ is formulated as

$$\left|\begin{array}{ll} \underset{\boldsymbol{x}, m}{\text{maximize}} & \displaystyle\sum_{s \in \mathcal{S}} p_s \, \mathcal{P}(\boldsymbol{x}, \boldsymbol{\xi}_s) - \lambda \left(m + \frac{1}{1-\beta} \sum_{s \in \mathcal{S}} p_s \, [-\mathcal{P}(\boldsymbol{x}, \boldsymbol{\xi}_s) - m]^+ \right) \\ \text{subject to} & \boldsymbol{x} \in X, \end{array}\right. \tag{3.22}$$

which is equivalent to the linear program (LP):

$$\left|\begin{array}{lll} \underset{\boldsymbol{x}, m, \boldsymbol{v}, \boldsymbol{w}, \boldsymbol{z}}{\text{maximize}} & \displaystyle\sum_{s \in \mathcal{S}} p_s \sum_{i \in \mathcal{N}} V_i \, \xi_{s,i} - \sum_{s \in \mathcal{S}} p_s \sum_{i \in \mathcal{N}} E_i \, w_{s,i} - \sum_{s \in \mathcal{S}} p_s \sum_{i \in \mathcal{N}} U_i \, z_{s,i} - \lambda \, m - \frac{\lambda}{1-\beta} \sum_{s \in \mathcal{S}} p_s \, v_s & \\ \text{subject to} & \displaystyle v_s \geq -\sum_{i \in \mathcal{N}} V_i \, \xi_{s,i} + \sum_{i \in \mathcal{N}} E_i \, w_{s,i} + \sum_{i \in \mathcal{N}} U_i \, z_{s,i} - m, \quad v_s \geq 0, & s \in \mathcal{S}, \\ & w_{s,i} \geq x_i - \xi_{s,i}, \quad w_{s,i} \geq 0, & s \in \mathcal{S}, \ \ i \in \mathcal{N}, \\ & z_{s,i} \geq \xi_{s,i} - x_i, \quad z_{s,i} \geq 0, & s \in \mathcal{S}, \ \ i \in \mathcal{N}, \\ & \boldsymbol{x} \in X. & \end{array}\right. \tag{3.23}$$

Proposition 3.8 *Let $(\boldsymbol{x}^*, m^*, \boldsymbol{v}^*, \boldsymbol{w}^*, \boldsymbol{z}^*)$ be an optimal solution of (3.23). Then, $(\boldsymbol{x}^*, m^*)$ is also optimal to (3.22), and the optimal objective values of both problems coincide.* ⋆

The minimization of the net loss CVaR with an expected profit constraint, i.e., (3.15) with $\rho(\mathcal{L}(\boldsymbol{x}, \tilde{\boldsymbol{\xi}})) = \mathrm{CVaR}_\beta(\boldsymbol{x})$, is

$$\left|\begin{array}{ll} \underset{\boldsymbol{x},m}{\text{minimize}} & m + \dfrac{1}{1-\beta} \displaystyle\sum_{s \in \mathcal{S}} p_s \left[-\mathcal{P}(\boldsymbol{x}, \boldsymbol{\xi}_s) - m \right]^+ \\ \text{subject to} & \displaystyle\sum_{s \in \mathcal{S}} p_s \, \mathcal{P}(\boldsymbol{x}, \boldsymbol{\xi}_s) \geq \overline{\mu}, \\ & \boldsymbol{x} \in X, \end{array}\right. \tag{3.24}$$

which is transformed to the following LP:

$$\left|\begin{array}{ll} \underset{\boldsymbol{x},m,\boldsymbol{v},\boldsymbol{w},\boldsymbol{z}}{\text{minimize}} & m + \dfrac{1}{1-\beta} \displaystyle\sum_{s \in \mathcal{S}} p_s \, v_s \\ \text{subject to} & \displaystyle\sum_{s \in \mathcal{S}} p_s \sum_{i \in \mathcal{N}} V_i \, \xi_{s,i} - \sum_{s \in \mathcal{S}} p_s \sum_{i \in \mathcal{N}} E_i \, w_{s,i} - \sum_{s \in \mathcal{S}} p_s \sum_{i \in \mathcal{N}} U_i \, z_{s,i} \geq \overline{\mu}, \\ & v_s \geq -\displaystyle\sum_{i \in \mathcal{N}} V_i \, \xi_{s,i} + \sum_{i \in \mathcal{N}} E_i \, w_{s,i} + \sum_{i \in \mathcal{N}} U_i \, z_{s,i} - m, \quad v_s \geq 0, \qquad s \in \mathcal{S}, \\ & w_{s,i} \geq x_i - \xi_{s,i}, \quad w_{s,i} \geq 0, \qquad s \in \mathcal{S}, \;\; i \in \mathcal{N}, \\ & z_{s,i} \geq \xi_{s,i} - x_i, \quad z_{s,i} \geq 0, \qquad s \in \mathcal{S}, \;\; i \in \mathcal{N}, \\ & \boldsymbol{x} \in X. \end{array}\right. \tag{3.25}$$

As readily seen, other variants using the total cost CVaR can also be transformed into equivalent LPs. In fact, only getting rid of the constant term, $-\sum_{i \in \mathcal{N}} V_i \, \xi_{s,i}$, from the constraint:

$$v_s \geq -\sum_{i \in \mathcal{N}} V_i \, \xi_{s,i} + \sum_{i \in \mathcal{N}} E_i \, w_{s,i} + \sum_{i \in \mathcal{N}} U_i \, z_{s,i} - m, \quad v_s \geq 0, \qquad s \in \mathcal{S}, \tag{3.26}$$

which is the first constraint of Problem (3.23) and the second constraint of Problem (3.25), we obtain two kinds of mean-risk models associated with the total cost CVaR. Such LP formulations have an overwhelming advantage especially when many constraints on multiple products need to be imposed. Besides, even when we cannot achieve any closed form solution, we can compute an (approximating) optimal solution.

In order to clarify the difference among the models discussed above and show the computational advantages of the LP formulations, several numerical results are presented below. We here consider the case of three products, i.e., $N = 3$, and $X = \mathbb{R}^3_+$. Scenarios of the demands for the products are generated from a multi-dimensional normal distribution, and p_s is set to be $1/S$ for every $s \in \mathcal{S}$. In the following experiments, we confirm that all the generated values $\xi_{s,i}$ become positive. We used Xpress-MP (ver.2005A) for Windows on a personal computer with Pentium4 processor (3.4GHz) and 2GB memory.

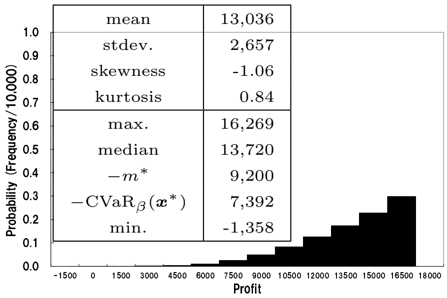
Figures 3.3 (a1) to (e2) show histograms and statistics of distributions of the profit $\mathcal{P}$ and the total cost $\mathcal{Q}$ via five models discussed above when ten thousand scenarios of their demand $(\tilde{\xi}_1, \tilde{\xi}_2, \tilde{\xi}_3)^\top$ are drawn from a three-dimensional normal distribution $\mathrm{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ where $\boldsymbol{\mu} \in \mathbb{R}^3$ with $\mu_i = 100$ for $i = 1, ..., 3$, and $\boldsymbol{\Sigma} := [\sigma^2_{ij}] \in \mathbb{R}^{3\times 3}$ with $\sigma^2_{ij} = 20^2$ $(i = j)$ and $0.5 \cdot 20^2$ $(i \neq j)$. In spite of the normality of the demand distribution, every distribution of the profit or the total cost is much skewed and, accordingly, far different from the normal one because of the nonlinearity of the profit function and the total cost functions.

We also observe some interesting differences among the proposed and classic models. From Figures 3.3 (a1) to (c2), we see that the resulting distributions through the classic profit maximization and the total cost CVaR minimization show relatively similar shapes, while the net loss CVaR minimization shows quite different distributions from those of the other two. In fact, the standard deviation of the profit $\mathcal{P}$ of the net loss CVaR minimizer is less than one third of that of the classic profit maximizer. In addition, the minimum value of the profit $\mathcal{P}$ of the net loss CVaR minimizer is positive, while those of the other two are negative. These facts encourage the use of the net loss CVaR measure as an alternative capturing the dispersion of the profit $\mathcal{P}$. In contrast to the profit $\mathcal{P}$, the distribution of the total cost $\mathcal{Q}$ by the net loss CVaR minimizer exhibits larger dispersion than the other two. From this fact, we should recognize that the definition of the loss function for the CVaR measure is crucial because the total cost CVaR minimizer and the net loss CVaR minimizer may result in having the totally different distributions of the profit $\mathcal{P}$ and the total cost $\mathcal{Q}$, as shown in these figures.

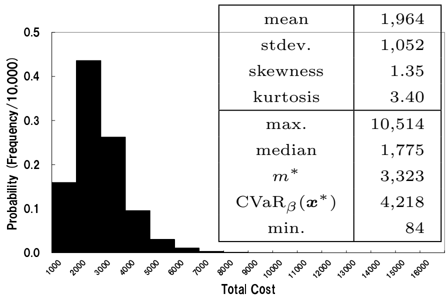
As for the mean-CVaR models, they can be used in tailoring a distribution to meet manager's demand. Indeed, as shown in Figures (d1) to (e2), it can be observed that each mean-CVaR model achieves a mixture of distributions via the mean profit maximization and a CVaR minimization of interest.

Figure 3.4 shows two kinds of convex efficient frontiers of the mean-net loss CVaR model with $\beta = 0.99$. Each line corresponds to an underlying normal distribution with different covariance matrix. One has uncorrelated covariance matrix, i.e., $\sigma_{ij} = 0.0$ for all $i \neq j$, while the other has positive covariance given by $\sigma_{ij}^2 = 0.5 \cdot 20^2$ for $i \neq j$. We here draw the efficient frontiers in an unusual manner so as to emphasize the change of the minimal CVaR with respect to that of the target return $\overline{\mu}$. Similarly to the mean-variance model [53], lower risk (i.e., lower CVaR) is attained in the case of lower correlation. Besides, the trade-off between the expected profit $\overline{\mu}$ and the minimal net loss CVaR is clearer in the case of higher correlation.

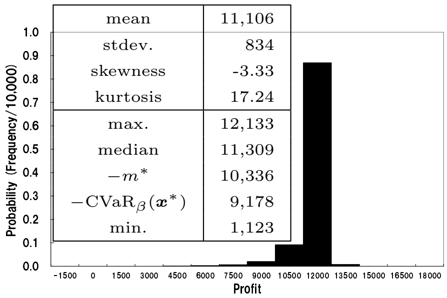
Finally, Figures 3.5 shows the average CPU time spent in solving the net loss CVaR minimization in LP form which is obtained from (3.25) by removing the first constraint. The average is taken over five sets of scenarios of size $S = 10{,}000$, 7,500 and 5,000. Although the computation time increases as β grows, all the problems are solved in reasonable time.



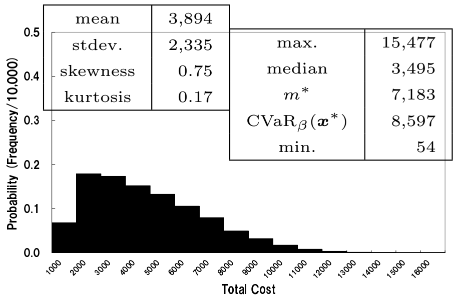
(a1) Profit by Classical Model



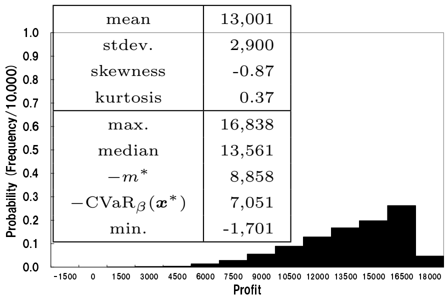
(a2) Total Cost by Classical Model



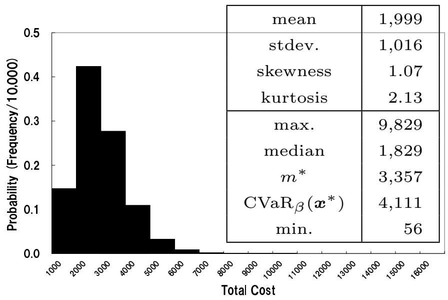
(b1) Profit by Net Loss CVaR Min.



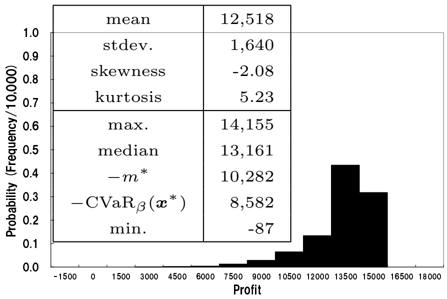
(b2) Total Cost by Net Loss CVaR Min.



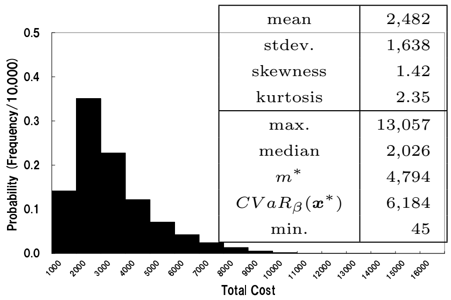
(c1) Profit by Total Cost CVaR Min.



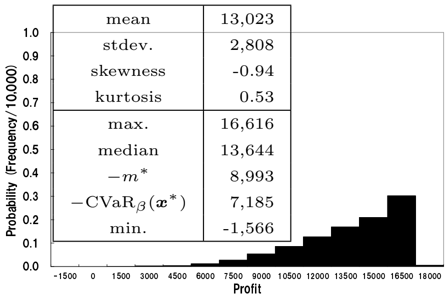
(c2) Total Cost by Total Cost CVaR Min.



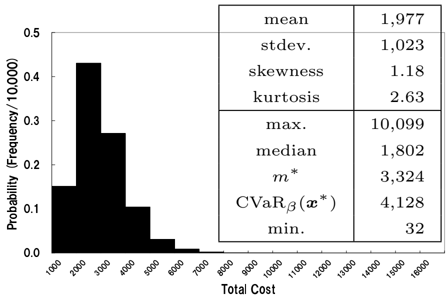
(d1) Profit by Mean-Net Loss CVaR ($\lambda = 1$)



(d2) Total Cost by Mean-Net Loss CVaR ($\lambda = 1$)



(e1) Profit by Mean-Total Cost CVaR ($\lambda = 0.5$)



(e2) Total Cost by Mean-Total Cost CVaR ($\lambda = 0.5$)

Figure 3.3: Histograms of Profit and Total Cost via Each Optimal Solution $(\boldsymbol{x}^*, m^*)$

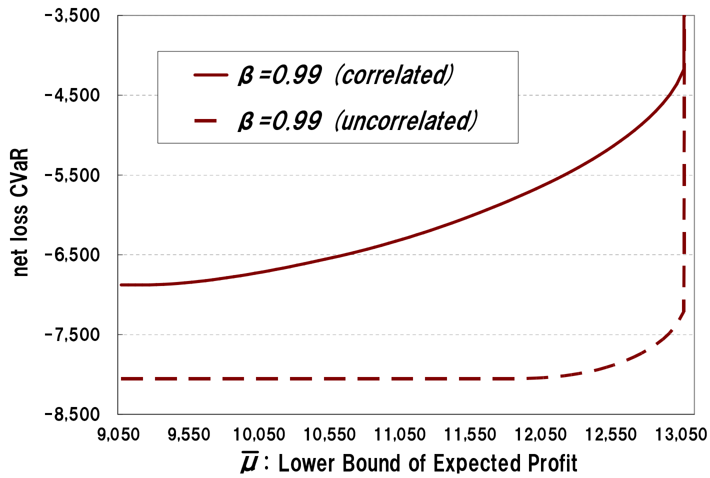


Figure 3.4: Efficient Frontiers of Mean-Net Loss CVaR Model under Two Distributions ($\beta = 0.99$)

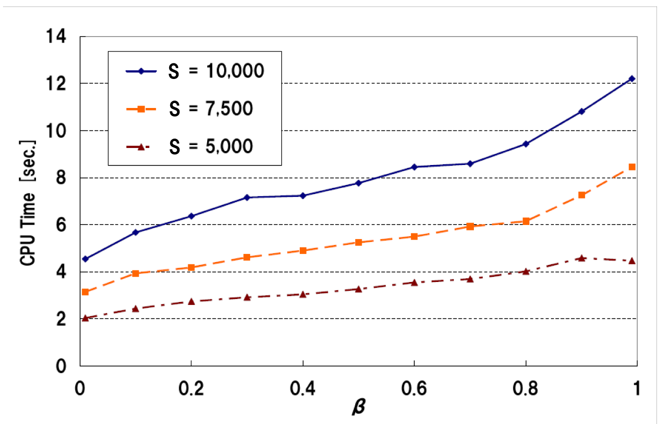


Figure 3.5: Average CPU Time for Solving LP Formulation and β [sec.]

Chapter 4

α-Conservative Approximation for Probabilistically Constrained Convex Programs

4.1 α-Conservative Approximation for PCCP

In order to tame a difficulty arising from the nonconvexity of PCCP (2.20), Nemirovski and Shapiro [63] introduce a convex conservative constraint (2.24), presenting a convex optimization problem which provides a feasible solution of the original problem (2.20). Although their approach enjoys the convex structure, the discrepancy from the original problem (2.20) is not clear. In this chapter, we extend the conservative approach by partly relinquishing convexity.

For a parameter $\alpha > 0$, let us define $\Psi_\alpha : \mathbb{R} \to \mathbb{R}$ by

$$\Psi_\alpha(z) := \Phi_{\alpha,1}(z) - \Phi_{\alpha,2}(z), \tag{4.1}$$

where

$$\Phi_{\alpha,1}(z) := \max\Big\{0,\, 1 + \frac{1}{\alpha}z\Big\}, \quad \Phi_{\alpha,2}(z) := \max\Big\{0,\, \frac{1}{\alpha}z\Big\}. \tag{4.2}$$

Similarly to the discussion about Nemirovski and Shapiro's convex conservative ap-

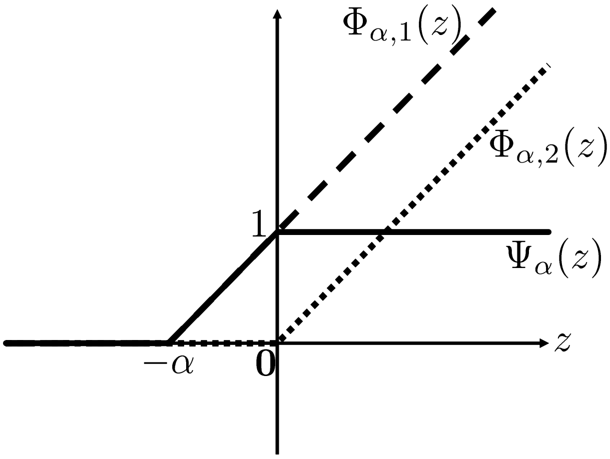


Figure 4.1: Graphs of Ψ_α, $\Phi_{\alpha,1}$ and $\Phi_{\alpha,2}$.

proximation in Section 2.2.2, we obtain a new conservative approximation problem:

$$
(\text{CAP}(\alpha)) \quad \left| \begin{array}{ll} \underset{\boldsymbol{x}\in\mathbb{R}^J}{\text{minimize}} & \hat{f}(\boldsymbol{x}) \\ \text{subject to} & \mathbb{E}\Big[\Psi_\alpha(g(\boldsymbol{x},\tilde{\boldsymbol{\xi}}))\Big] \le 1-\beta \\ & \boldsymbol{x} \in X. \end{array} \right. \tag{4.3}
$$

We refer to this problem as *α-conservative approximation* problem of (2.20), and the new constraint as *α-conservative approximation constraint.* Note that both $\mathbb{E}\Big[\Phi_{\alpha,1}(g(\boldsymbol{x},\tilde{\boldsymbol{\xi}}))\Big]$ and $\mathbb{E}\Big[\Phi_{\alpha,2}(g(\boldsymbol{x},\tilde{\boldsymbol{\xi}}))\Big]$ are convex in $\boldsymbol{x}$ since both $\Phi_{\alpha,1}$ and $\Phi_{\alpha,2}$ are nondecreasing convex functions. Accordingly,

$$
\mathbb{E}\Big[\Psi_\alpha(g(\boldsymbol{x},\tilde{\boldsymbol{\xi}}))\Big] = \mathbb{E}\Big[\Phi_{\alpha,1}(g(\boldsymbol{x},\tilde{\boldsymbol{\xi}}))\Big] - \mathbb{E}\Big[\Phi_{\alpha,2}(g(\boldsymbol{x},\tilde{\boldsymbol{\xi}}))\Big]
$$

is a D.C. function and Problem (4.3) is a D.C. optimization problem, for which several global optimization algorithms have been developed (e.g. Tuy [81]).

It is easy to show that the new constraint approaches the original probabilistic constraint as α decreases to 0 in the following sense.

Proposition 4.1 $\mathbb{E}\Big[\Psi_\alpha(g(\boldsymbol{x},\tilde{\boldsymbol{\xi}}))\Big] - \text{VP}(\boldsymbol{x}) \to \text{Prob}\Big\{g(\boldsymbol{x},\tilde{\boldsymbol{\xi}}) = 0\Big\}$ *for each* $\boldsymbol{x}$, *as* $\alpha \to +0$. *Especially, if* $g(\boldsymbol{x},\tilde{\boldsymbol{\xi}})$ *has a continuous cumulative distribution function, one then has* $\mathbb{E}\Big[\Psi_\alpha(g(\boldsymbol{x},\tilde{\boldsymbol{\xi}}))\Big] \to \text{VP}(\boldsymbol{x})$ *for each* $\boldsymbol{x}$, *as* $\alpha \to +0$. ■

From this proposition, solving CAP(α) for a small α yields a good approximate solution of the original problem (2.20).

In general, $\mathbb{E}\Big[\Psi_\alpha(g(\boldsymbol{x},\tilde{\boldsymbol{\xi}}))\Big]$ is nondifferentiable at some points because of the nondifferentiability of $\Phi_{\alpha,1}$ and $\Phi_{\alpha,2}$. By employing differentiable functions in place of $\Phi_{\alpha,1}$ and $\Phi_{\alpha,2}$ in Problem (4.3), one can construct differentiable α-conservative approximations even for discrete distribution.

Example of Differentiable α-Conservative Constraint The α-conservative approximation of the probabilistic constraint can be constructed in other ways. The key of the construction is approximating the max function in (4.2) by a smoothing function with a parameter $\delta > 0$ ([66]). Two examples of such a smoothing function are

$$\rho_{\delta,1}(z) := \frac{\sqrt{z^2+4\delta^2}+z}{2}, \qquad \rho_{\delta,2}(z) := \delta\log(1+\exp(z/\delta)).$$

Proposition 4.2 *If ϕ is a convex, nondecreasing function, then $\psi_\alpha(z) := \phi(1+z/\alpha) - \phi(z/\alpha)$ is a D.C. function which is nonnegative valued, nondecreasing and satisfies*

$$\psi_\alpha(0)\mathbb{1}_{[0,+\infty)}(z) \le \psi_\alpha(z) \quad \text{for all } z \in \mathbb{R}. \tag{4.4}$$

⋆

Proof : It is clear that ψ_α is nonnegative valued and nondecreasing. Then, (4.4) holds because ψ_α is nonnegative valued for $z < 0$, and because ψ_α is nondecreasing for $z \ge 0$. ■

Substituting $g(\boldsymbol{x},\tilde{\boldsymbol{\xi}})$ for z and taking mathematical expectation in (4.4), we have

$$\mathbb{E}\Big[\psi_\alpha(g(\boldsymbol{x},\tilde{\boldsymbol{\xi}}))\Big] \le \psi_\alpha(0)(1-\beta) \;\Rightarrow\; \mathrm{VP}(\boldsymbol{x}) \le 1-\beta.$$

By adopting the above smoothing functions $\rho_{\delta,1}, \rho_{\delta,2}$ as a function ϕ in Proposition 4.2, different α-conservative approximations are derived. The D.C. functions thus obtained are differentiable, and the basic framework of the algorithms which we will describe in Section 4.3 can be applied to those formulations. In the following sections, our discussion focuses on the D.C. inequality composed of the piecewise linear function (4.1).

4.2 Portfolio Selection via Value-at-Risk Minimization

In this section, we formulate the minimization of the Value-at-Risk (VaR) of a financial asset portfolio as an example of the PCCP.

The VaR minimization of a financial asset portfolio is to determine the amount of investment (or investment ratio) to N kinds of financial assets so that it achieves the minimum β-VaR. Formally, it is formulated as the following optimization problem:

$$\left|\begin{array}{ll} \underset{(\boldsymbol{w},m)\in\mathbb{R}^N\times\mathbb{R}}{\text{minimize}} & m \\ \text{subject to} & \text{Prob}\Big\{\boldsymbol{w}^\top\tilde{\boldsymbol{y}} - m > 0\Big\} \leq 1-\beta \\ & \boldsymbol{w}\in W, \end{array}\right. \tag{4.5}$$

where

$\boldsymbol{w}$: investment ratio to N kinds of financial assets (decision variable), $\boldsymbol{w}\in\mathbb{R}^N$

m : VaR (decision variable), $m\in\mathbb{R}$

W : set of feasible portfolio $\boldsymbol{w}$, $W\subseteq\mathbb{R}^N$

$\tilde{\boldsymbol{y}}$: N dimensional random vector representing the loss associated with the financial assets

β : confidence level, $\beta\in(0,1)$.

The feasible set W is defined by several constraints such as the minimal expected return constraint $\bar{\boldsymbol{r}}^\top\boldsymbol{w}\geq R$ where $\bar{\boldsymbol{r}}$ is the N dimensional expected return vector of the financial assets, and R is the the minimal expected return of the portfolio. The random loss $\tilde{\boldsymbol{y}}$ is sometimes defined as "$(-1)\times$(rate of return)".

In the rest of this section, we denote by $\mathcal{S}=\{1,...,S\}$ a index set of scenarios and assume that

Assumption 4.3 *We have a finite set of scenarios $\{\boldsymbol{y}^s \mid s\in\mathcal{S}\}$ of the random loss $\tilde{\boldsymbol{y}}$.* ⋆

We denote by p_s the occurrence probability of scenario $\boldsymbol{y}^s$ and assume that $\sum\limits_{s\in\mathcal{S}} p_s = 1$ and $p_s>0$ for all $s\in\mathcal{S}$.

Furthermore, we assume the following:

Assumption 4.4 *The feasible region W of $\boldsymbol{w}$ is a polytope.* $\star$

This assumption seems reasonable since the constraints $\mathbf{1}^\top \boldsymbol{w} = 1$ and $\boldsymbol{w} \geq \mathbf{0}$ are included in many practical situations. In addition, the minimum return constraint, upper limit of investment ratio and the like are representable by linear inequalities.

The most typical way to an exact solution of Problem (4.5) is reformulating it as a 0-1 mixed integer program:

$$\left|\begin{array}{ll} \underset{(\boldsymbol{w},m,\boldsymbol{u})\in\mathbb{R}^N\times\mathbb{R}\times\mathbb{R}^S}{\text{minimize}} & m \\ \text{subject to} & \displaystyle\sum_{s\in\mathcal{S}} p_s u_s \leq 1-\beta \\ & \boldsymbol{w}^\top \boldsymbol{y}^s - m \leq U u_s,\ u_s \in \{0,1\}, \quad s \in \mathcal{S} \\ & \boldsymbol{w} \in W, \end{array}\right. \tag{4.6}$$

where U is a sufficiently large number satisfying

$$U > \max\{\boldsymbol{w}^\top \boldsymbol{y}^s \,|\, \boldsymbol{w} \in W,\ s \in \mathcal{S}\} - \min\{\boldsymbol{w}^\top \boldsymbol{y}^s \,|\, \boldsymbol{w} \in W,\ s \in \mathcal{S}\}.$$

Note that Problem (4.6) has 0-1 variables as many as scenarios. This would be a disadvantage of this formulation when we consider a large number of scenarios to enhance the reliability of the solution to be obtained.

In the following sections, we consider the α-conservative approximation of Problem (4.5) under Assumptions 4.3 and 4.4, as follows:

$$\left|\begin{array}{ll} \underset{(\boldsymbol{w},m)\in\mathbb{R}^N\times\mathbb{R}}{\text{minimize}} & m \\ \text{subject to} & \sum_{s\in\mathcal{S}} p_s \Phi_{\alpha,1}(\boldsymbol{w}^\top \boldsymbol{y}^s - m) - \sum_{s\in\mathcal{S}} p_s \Phi_{\alpha,2}(\boldsymbol{w}^\top \boldsymbol{y}^s - m) \leq 1-\beta \\ & \boldsymbol{w} \in W. \end{array}\right. \tag{4.7}$$

4.3 Global Optimization Algorithm

In this section, a simplicial branch-and-bound algorithm is presented for computing a globally optimal solution of Problem (4.7). Some remarks on the application of an outer approximation algorithm for Problem (4.7) will be provided.

4.3.1 Simplicial Branch-and-Bound Algorithm

By denoting

$$h^D(\boldsymbol{w}, m) := \sum_{s\in\mathcal{S}} p_s \Phi_{\alpha,1}(\boldsymbol{w}^\top \boldsymbol{y}^s - m), \quad h^C(\boldsymbol{w}, m) := \sum_{s\in\mathcal{S}} p_s \Phi_{\alpha,2}(\boldsymbol{w}^\top \boldsymbol{y}^s - m),$$

Problem (4.7) is rewritten as

$$\left| \begin{array}{ll} \underset{(\boldsymbol{w},m)\in\mathbb{R}^N\times\mathbb{R}}{\text{minimize}} & m \\ \text{subject to} & h^D(\boldsymbol{w}, m) - h^C(\boldsymbol{w}, m) \le 1-\beta \\ & \boldsymbol{w} \in W. \end{array} \right. \tag{4.8}$$

Let $M \subset \mathbb{R}^{N+1}$ be the simplex being the convex hull of affinely independent vertices $\boldsymbol{v}^{M,1}, \boldsymbol{v}^{M,2}, ..., \boldsymbol{v}^{M,N+2}$. For M, we define the relaxed subproblem, RSP(M) of Problem (4.8):

$$(\mathrm{RSP}(M)) \left| \begin{array}{ll} \underset{\boldsymbol{\lambda}\in\mathbb{R}^{N+2}}{\text{minimize}} & \sum_{i=1}^{N+2} \lambda_i v^{M,i}_{N+1} \\ \text{subject to} & h^D\Big(\sum_{i=1}^{N+2} \lambda_i \boldsymbol{v}^{M,i}\Big) - \sum_{i=1}^{N+2} \lambda_i h^C(\boldsymbol{v}^{M,i}) \le 1-\beta \\ & \sum_{i=1}^{N+2} \lambda_i \boldsymbol{v}^{M,i} \in W \times [m_L,\, m_U],\ \boldsymbol{\lambda} \ge \mathbf{0},\ \mathbf{1}^\top\boldsymbol{\lambda} = 1, \end{array} \right. \tag{4.9}$$

where m_L and m_U are, respectively, lower and upper bounds on the optimal objective value of Problem (4.8).

It is easy to see that RSP(M) provides a lower bound of the objective value of Problem (4.8) over M. Technically, m_L can be computed via an algorithm of Pang and Leyffer [66], for example, and m of any feasible solution $(\boldsymbol{w}, m)$ can be employed as m_U, whereas, we used sufficiently small and large numbers as m_L and m_U, respectively in the experiments reported in Section 4.4.

The initial simplex M_0 contains $W \times [m_L,\, m_U]$ so that an optimal solution of Problem (4.8) is contained in M_0. For such M_0, we solve RSP(M_0), obtaining a lower bound on the optimal value of Problem (4.8). Note that one can easily find a feasible solution of Problem (4.8) if $\boldsymbol{w} \in W$ is available because a sufficiently large m satisfies the D.C. inequality. In addition, due to the monotonicity of the left-hand side of the D.C. inequality with respect to m, we can find m satisfying the inequality by equality. Such an m serves as the incumbent value (i.e., the best known upper bound on the optimal value).

In the algorithm we propose, as long as any simplex remains to be considered, we choose a simplex M with the lowest lower bound. We first bisect the simplex M i.e., split M at the middle point of the longest edge into two simplices M' and M'', and replace M with two simplices M' and M'' *(branching procedure).* Although other simplex splitting rules such as $\boldsymbol{\omega}$-subdivision are also suggested, we use the exact bisection rule because it is the simplest one and guarantees that the sequence of simplices shrinks to a single point. We then compute lower bounds over M' and M'' by solving RSP(M') and RSP(M''), respectively. If we find a feasible solution of Problem (4.8) with objective value smaller than the incumbent in the solution process of RSP(M') and RSP(M''), the incumbent is updated with the better solution. Let γ be the incumbent objective value, i.e., the best objective value obtained so far. If the lower bound on a simplex M is no less than γ, we discard it from further consideration *(bounding procedure).* If there is no simplex to be considered, the algorithm terminates and the global optimality is guaranteed.

A Procedure for Searching Good Incumbents In order to improve the incumbent solution, it is better to obtain a feasible solution of Problem (4.8) whenever RSP(M) returns a solution $(\boldsymbol{w}(M), m(M))$. However, $(\boldsymbol{w}(M), m(M))$ is rarely feasible to Problem (4.8). Since $\boldsymbol{w}(M)$ belongs to W at all times, and the left-hand side of the D.C. constraint of Problem (4.8) is nonincreasing with respect to m, we can construct a procedure for finding a feasible solution to Problem (4.8) from an infeasible solution $(\boldsymbol{w}(M), m(M))$ and replacing the incumbent with the solution if it achieves lower objective value.

Subroutine for searching a feasible solution from an infeasible solution
Let $(\boldsymbol{w}(M), m(M))$ be infeasible to Problem (4.8) and $\boldsymbol{w}(M) \in W$, and let γ be the incumbent objective value. Check if $(\boldsymbol{w}(M), \gamma)$ satisfies $h^D(\boldsymbol{w}(M), \gamma) - h^C(\boldsymbol{w}(M), \gamma) \leq 1 - \beta$.

(a) If it holds true, search $\hat{m}$ satisfying $h^D(\boldsymbol{w}(M), \hat{m}) - h^C(\boldsymbol{w}(M), \hat{m}) = 1 - \beta$. (In this case, $\hat{m}$ is no more than γ and can be a new incumbent.)

(b) Otherwise, quit this subroutine.

By applying this procedure to each solution of RSP(M), we anticipate obtaining a good incumbent solution at an early stage of the branch-and-bound algorithm. This procedure reduces the size of branch-and-bound tree, and improves the performance of the algorithm. As already mentioned, an initial incumbent solution in the following algorithm is easily obtained since this subroutine can be used for any $\boldsymbol{w} \in W$.

We are now in a position to describe the simplicial branch-and-bound algorithm.

Algorithm SBB: Simplicial Branch-and-Bound Algorithm

Step 0. [**Initialization**] : Let $(\bar{\boldsymbol{w}}^1, \bar{m}^1)$ be an incumbent solution, M_0 be the initial simplex and ε (≥ 0) be a tolerance for optimality. Set $\mathcal{P}_1 \leftarrow \{M_0\}$, $\mathcal{Q}_1 \leftarrow \{M_0\}$, $\gamma^1 \leftarrow \bar{m}^1$ and $k \leftarrow 1$.

Step 1. [**Lower Bound Computation**] : For each $M \in \mathcal{P}_k$, let $(\boldsymbol{w}(M), m(M))$ be an optimal solution of RSP(M), where $m(M) \leftarrow +\infty$ if it is infeasible.

Step 2. [**Incumbent Solution Update**] : Let $(\bar{\boldsymbol{w}}, \bar{m}) \leftarrow (\bar{\boldsymbol{w}}^k, \bar{m}^k)$, $\gamma \leftarrow \gamma^k$.
for all $M \in \mathcal{P}_k$
  if $m(M) < \gamma$
    if $(\boldsymbol{w}(M), m(M))$ is feasible to Problem (4.8)
      $(\bar{\boldsymbol{w}}, \bar{m}) \leftarrow (\boldsymbol{w}(M), m(M))$, $\gamma \leftarrow m(M)$
    else
      find a feasible solution $(\boldsymbol{w}(M), \hat{m}(M))$ to Problem (4.8)
      via the subroutine above

if a feasible solution $(\boldsymbol{w}(M), \hat{m}(M))$ such that $\hat{m}(M) < \gamma$ is found

$(\bar{\boldsymbol{w}}, \bar{m}) \leftarrow (\boldsymbol{w}(M), \hat{m}(M)),\ \gamma \leftarrow \hat{m}(M)$

end if

end if

end if

end for

Set $(\bar{\boldsymbol{w}}^{k+1}, \bar{m}^{k+1}) \leftarrow (\bar{\boldsymbol{w}}, \bar{m}),\ \gamma^{k+1} \leftarrow \gamma$.

Step 3. [Bounding] : Set $\mathcal{R}_k \leftarrow \{M \in \mathcal{Q}_k \mid m(M) < \gamma^{k+1} - \varepsilon\}$.

Step 4. [Optimality Check] : If $\mathcal{R}_k = \emptyset$, terminate the algorithm with $(\bar{\boldsymbol{w}}^{k+1}, \bar{m}^{k+1})$ as an ε-optimal solution of Problem (4.8).

Step 5. [Branching] : Select $M^* \in \arg\min\{m(M) \mid M \in \mathcal{R}_k\}$ and bisect M^* into two simplices. Let $\mathcal{P}_{k+1}$ be the simplices then obtained.

Step 6. [Simplex Set Update] : Set $\mathcal{Q}_{k+1} \leftarrow (\mathcal{R}_k \setminus \{M^*\}) \cup \mathcal{P}_{k+1},\ \ k \leftarrow k+1$ and return to Step 1.

Proposition 4.5 *If $\varepsilon > 0$, the above algorithm terminates after finitely many iterations. If $\varepsilon = 0$, the above algorithm can repeat infinitely, and in this case, every accumulation point of the sequence of incumbent solutions $\{(\bar{\boldsymbol{w}}^k, \bar{m}^k)\}$ is a globally optimal solution to Problem (4.8).* ⋆

Proof : This proposition follows Proposition 5.6 in Tuy [81]. ■

4.3.2 Computation of the Relaxed Problems

In the above branch-and-bound algorithm, each relaxed problem RSP(M) on a simplex M is a convex program with a single nonlinear constraint $h^D(\sum_{i=1}^{N+2} \lambda_i \boldsymbol{v}^{M,i}) - \sum_{i=1}^{N+2} \lambda_i h^C(\boldsymbol{v}^{M,i}) \leq 1 - \beta$. Since $h^D(\sum_{i=1}^{N+2} \lambda_i \boldsymbol{v}^{M,i})$ is a convex and piecewise linear function in $\boldsymbol{\lambda}$, RSP(M) could be formulated as a linear program by introducing additional decision variables and constraints. However, this LP formulation is not appropriate for VaR minimization problem because the numbers of decision variables and constraints would be heavily dependent on the number of scenarios, which is usually enormous. We here consider the following three different implementation strategies for solving the relaxed problem:

1) Direct application of a nonlinear optimization solver to the problem

2) Relaxation of the convex function h^D by selecting a part of linear functions which constitute h^D, and obtaining a linear programming formulation.

3) Application of Kelley's cutting plane method in which a sequence of LPs is iteratively solved.

The first strategy makes the number of variables and constraints independent of the number of scenarios, and therefore, it seems to exploit the preferable characteristics of RSP(M). However, in applying a nonlinear optimization solver to solving RSP(M), we should develop a subroutine for the infeasibility check, which reduces the computational efficiency.

To overcome this drawback, we employ LP based subroutines for computing the lower bound. The first alternative strategy uses a part of linear functions which coincides with h^D at extreme points and the center of each simplex. We solve a relaxed problem of RSP(M) in this strategy while the size of the resulting LP is still independent of the number of scenarios.

Another alternative is a straightforward application of the well-known Kelley's cutting plane method. The relaxed problem RSP(M) is solved in an exact manner in this strategy, and accordingly, it may deal with a number of constraints as many as scenarios. However, this strategy is expected to work efficiently because it brings in the constraints effectively when needed, and the efficient dual simplex algorithm can be applied.

The details of the second and third strategies are described as follows:

Linear relaxation for the relaxed subproblem In building a linear relaxation LR(M) for the relaxed subproblem RSP(M), we use a linear approximation of the function h^D at the extreme points and the center of M. In the branch-and-bound algorithm, we solve LR(M) on a simplex M in place of RSP(M).

$$
(\mathrm{LR}(M)) \left|
\begin{array}{ll}
\underset{\boldsymbol{\lambda} \in \mathbb{R}^{N+2}}{\text{minimize}} & \displaystyle\sum_{i=1}^{N+2} \lambda_i v_{N+1}^{M,i} \\
\text{subject to} & h^D\Big(\displaystyle\sum_{i=1}^{N+2} \frac{\boldsymbol{v}^{M,i}}{N+2}\Big) + \Big(\boldsymbol{V}_M \boldsymbol{h}_{sg}^D\Big(\sum_{i=1}^{N+2} \frac{\boldsymbol{v}^{M,i}}{N+2}\Big)\Big)^\top \Big(\boldsymbol{\lambda} - \frac{\mathbf{1}}{N+2}\Big) \\
& \qquad - \displaystyle\sum_{i=1}^{N+2} \lambda_i h^C(\boldsymbol{v}^{M,i}) \le 1 - \beta \\
& h^D(\boldsymbol{v}^{M,j}) + \Big(\boldsymbol{V}_M \boldsymbol{h}_{sg}^D(\boldsymbol{v}^{M,j})\Big)^\top (\boldsymbol{\lambda} - \boldsymbol{e}^j) \\
& \qquad - \displaystyle\sum_{i=1}^{N+2} \lambda_i h^C(\boldsymbol{v}^{M,i}) \le 1 - \beta, \qquad j = 1, ..., N+2 \\
& \displaystyle\sum_{i=1}^{N+2} \lambda_i \boldsymbol{v}^{M,i} \in W \times [m_L,\, m_U],\ \boldsymbol{\lambda} \ge \mathbf{0},\ \mathbf{1}^\top \boldsymbol{\lambda} = 1,
\end{array}
\right.
\tag{4.10}
$$

where $\boldsymbol{e}^j$ is the j-th unit vector for $j = 1, ..., N+2$,

$$
\boldsymbol{V}_M := \Big(\boldsymbol{v}^{M,1}\ \boldsymbol{v}^{M,2}\ \cdots\ \boldsymbol{v}^{M,N+2}\Big)^\top,
$$

and $\boldsymbol{h}_{sg}^D(\bar{\boldsymbol{v}})$ is a subgradient of $h^D(\boldsymbol{v})$ at $\boldsymbol{v} = \bar{\boldsymbol{v}}$.

Remark 4.6 When the number of scenarios is very large, it is heavy to calculate the function values of h^D, $\boldsymbol{h}_{sg}^D$ and h^C. However, we can reduce such a computational burden in the above strategy by reusing calculated function values. ⋆

Kelley's cutting plane method for the relaxed subproblem Let $\hat{\boldsymbol{\lambda}}^1$ be a solution of LR(M). Note that if LR(M) is infeasible, so is RSP(M). Let κ (> 0) be a sufficiently small number. If $h^D\Big(\sum_{i=1}^{N+2} \hat{\lambda}_i^1 \boldsymbol{v}^{M,i}\Big) - \sum_{i=1}^{N+2} \hat{\lambda}_i^1 h^C(\boldsymbol{v}^{M,i}) \le 1 - \beta + \kappa$ is not satisfied, we add a constraint

$$
h^D\Big(\sum_{i=1}^{N+2} \hat{\lambda}_i^1 \boldsymbol{v}^{M,i}\Big) + \Big((\boldsymbol{V}_M)^\top \boldsymbol{h}_{sg}^D\Big(\sum_{i=1}^{N+2} \hat{\lambda}_i^1 \boldsymbol{v}^{M,i}\Big)\Big)^\top (\boldsymbol{\lambda} - \hat{\boldsymbol{\lambda}}^1) - \sum_{i=1}^{N+2} \hat{\lambda}_i^1 h^C(\boldsymbol{v}^{M,i}) \le 1 - \beta,
$$

to LR(M), and let $\hat{\boldsymbol{\lambda}}^2$ be a solution of the augmented problem. If $h^D\Big(\sum_{i=1}^{N+2} \hat{\lambda}_i^2 \boldsymbol{v}^{M,i}\Big) - \sum_{i=1}^{N+2} \hat{\lambda}_i^2 h^C(\boldsymbol{v}^{M,i}) \le 1 - \beta + \kappa$ is not satisfied, add the constraint on $\hat{\boldsymbol{\lambda}}^2$ and solve the new problem. Since h^D and h^C are piecewise linear functions, this procedure will end up with an optimal solution of RSP(M) within a finite number of iterations. We, however, set the upper limit on the number of iterations for the sake of overall efficiency.

4.3.3 Storage of VaR Best Solutions

Let VaR($\boldsymbol{w}$) denote the VaR of a portfolio $\boldsymbol{w}$. VaR($\boldsymbol{w}$) is evaluated as follows:

1. Sort $\{\boldsymbol{w}^\top \boldsymbol{y}^s\}_{s\in\mathcal{S}}$ in descending order as $\boldsymbol{w}^\top \boldsymbol{y}^{s_1} \geq ... \geq \boldsymbol{w}^\top \boldsymbol{y}^{s_S}$.
2. For $\hat{j} := \min\{\,\theta\,:\, \sum_{j=1}^{\theta} p_{s_j} > 1-\beta\,\}$, $\mathrm{VaR}(\boldsymbol{w}) = \boldsymbol{w}^\top \boldsymbol{y}^{s_{\hat{j}}}$.

We should notice that VaR($\bar{\boldsymbol{w}}^k$) is not always equal to $\bar{m}^k$ for an incumbent solution $(\bar{\boldsymbol{w}}^k, \bar{m}^k)$. Therefore, it may occur that $(\bar{\boldsymbol{w}}^k, \bar{m}^k)$ is updated by $(\bar{\boldsymbol{w}}^{k+1}, \bar{m}^{k+1})$ whereas VaR($\bar{\boldsymbol{w}}^k$) is no more than VaR($\bar{\boldsymbol{w}}^{k+1}$). We refer to the solution selected according to VaR($\boldsymbol{w}$) in place of m as VaR best solution. In order to obtain a better estimation of the minimal VaR, we employ the following strategy.

Add "set $(\bar{\boldsymbol{w}}^1_{\mathrm{VaR}}, \bar{m}^1_{\mathrm{VaR}}) \leftarrow (\bar{\boldsymbol{w}}^1, \bar{m}^1)$" in Step 0 of the algorithm SBB and
"**if** $\mathrm{VaR}(\bar{\boldsymbol{w}}) < \mathrm{VaR}(\bar{\boldsymbol{w}}^k_{\mathrm{VaR}})$
$(\bar{\boldsymbol{w}}^{k+1}_{\mathrm{VaR}}, \bar{m}^{k+1}_{\mathrm{VaR}}) \leftarrow (\bar{\boldsymbol{w}}, \bar{m})$
else
$(\bar{\boldsymbol{w}}^{k+1}_{\mathrm{VaR}}, \bar{m}^{k+1}_{\mathrm{VaR}}) \leftarrow (\bar{\boldsymbol{w}}^k_{\mathrm{VaR}}, \bar{m}^k_{\mathrm{VaR}})$
end if"
at the end of Step 2, and replace
"terminate the algorithm with $(\bar{\boldsymbol{w}}^{k+1}, \bar{m}^{k+1})$ as an ε-optimal solution of Problem (4.8)."
with
"terminate the algorithm with $(\bar{\boldsymbol{w}}^{k+1}_{\mathrm{VaR}}, \bar{m}^{k+1}_{\mathrm{VaR}})$ as a VaR best solution of Problem (4.8)."
in Step 4.

Note that the behavior of the algorithm remains the same as the original version except the solution provided when the algorithm terminates. Though the sequence of VaR best solutions $\{(\bar{\boldsymbol{w}}^k_{\mathrm{VaR}}, \bar{m}^k_{\mathrm{VaR}})\}$ may not converge to an optimal solution of Problem (4.8), the VaR best solution is expected to have smaller VaR than that without the strategy. Moreover, the VaR best solution is more effective when the algorithm quits before satisfying the optimality.

Remark 4.7 *(On the Application of the Outer Approximation Method)* The second approach to solve the D.C. problem (4.7) is an outer approximation algorithm. By introducing a new variable π, Problem (4.7) is rewritten as follows:

$$\left|\begin{array}{ll} \underset{(\boldsymbol{w},m,\pi)\in\mathbb{R}^N\times\mathbb{R}\times\mathbb{R}}{\text{minimize}} & m \\ \text{subject to} & \sum_{s\in\mathcal{S}} p_s\Phi_{\alpha,1}(\boldsymbol{w}^\top\boldsymbol{y}^s - m) - \pi \le 1-\beta \\ & \sum_{s\in\mathcal{S}} p_s\Phi_{\alpha,2}(\boldsymbol{w}^\top\boldsymbol{y}^s - m) - \pi \ge 0 \\ & \boldsymbol{w}\in W. \end{array}\right. \tag{4.11}$$

By introducing two sets in $\mathbb{R}^{N+2}$ defined by

$$D := \left\{(\boldsymbol{w},m,\pi)\,\middle|\, g^D(\boldsymbol{w},m,\pi)\le 0,\ \boldsymbol{w}\in W\right\},\quad C := \left\{(\boldsymbol{w},m,\pi)\,\middle|\, g^C(\boldsymbol{w},m,\pi)\le 0\right\},$$

where $g^D(\boldsymbol{w},m,\pi) := h^D(\boldsymbol{w},m) - \pi - (1-\beta)$ and $g^C(\boldsymbol{w},m,\pi) := h^C(\boldsymbol{w},m) - \pi$, Problem (4.11) is reformulated as the following D.C. program:

$$\left|\ \underset{(\boldsymbol{w},m,\pi)\in\mathbb{R}^N\times\mathbb{R}\times\mathbb{R}}{\text{minimize}}\quad m \quad \text{subject to}\quad (\boldsymbol{w},m,\pi)\in D\setminus \text{int}C.\right. \tag{4.12}$$

We apply the outer approximation method in [81] to (4.12). Through some preliminary computational experiment, this method is found to be inferior to the simplicial branch-and-bound method combined with several strategies, and therefore, experimental result of this method will be omitted in this article. ⋆

4.4 Computational Experiments

In this section, we report some numerical results of the VaR minimization algorithms. We consider the minimization of the VaR of a portfolio consisting of five or ten assets, i.e., $N = 5$ or 10. The loss $\tilde{y}_i$ of asset i is given as an independent random variable.

The problem is

$$\left|\begin{array}{ll} \underset{(\boldsymbol{w},m)\in\mathbb{R}^N\times\mathbb{R}}{\text{minimize}} & m \\ \text{subject to} & \text{Prob}\left\{\sum_{i=1}^{N} w_i\tilde{y}_i - m > 0\right\} \leq 0.1 \\ & 0 \leq w_i \leq 0.49, \quad i = 1, ..., N \\ & \sum_{i=1}^{N} w_i = 1, \ \sum_{i=1}^{N} \bar{r}_i w_i \geq 1.2. \end{array}\right. \tag{4.13}$$

We set $\bar{r}_i = 1.25$ when i is odd and $\bar{r}_i = 1.1$ when i is even. The loss scenarios of assets 1, 2, 6 and 7 are drawn from a Cauchy distribution where the location of the peak of the density is 0, and the half-width at half-maximum is 2. The loss scenarios of assets 3, 4, 5, 8, 9 and 10 are drawn from a uniform distribution on the interval $[-12.5, 12.5]$. We consider three different number of scenarios 100, 1,000, 10,000, and set $p_s := \frac{1}{S}$ for all $s \in \mathcal{S}$.

We implemented five approaches to a solution of Problem (4.13): (a) the proposed branch-and-bound algorithm with linear relaxation, (b) the proposed branch-and-bound algorithm with Kelley's method, (c) the convex approximation (2.25) by Nemirovski and Shapiro [63] using $\psi(z) = \max\{0, 1+z\}$, (d) the CVaR minimization, (e) the typical MIP formulation (4.6), and we compared these in terms of the resulting $\text{VaR}(\boldsymbol{w}^*)$ and the violation probability $\text{VP}(\boldsymbol{w}^*, m^*) := \text{Prob}\{(\boldsymbol{w}^*)^\top \tilde{\boldsymbol{y}} - m^* > 0\}$ of the obtained solution $(\boldsymbol{w}^*, m^*)$. (a) and (b) are the proposed simplicial branch-and-bound algorithms with the storage of VaR best solution, and solve the relaxed subproblem by the two relaxation strategies, and we set $\varepsilon = 0.5$ as the tolerance for optimality. (d) is the Conditional Value-at-Risk (CVaR) minimization formulated as the following LP ([74]):

$$\left|\begin{array}{ll} \underset{(\boldsymbol{w},m,\boldsymbol{\tau})\in\mathbb{R}^N\times\mathbb{R}\times\mathbb{R}^S}{\text{minimize}} & m + \dfrac{1}{1-\beta}\sum_{s\in\mathcal{S}} p_s\tau_s \\ \text{subject to} & \tau_s \geq 0, \ \tau_s \geq \boldsymbol{w}^\top \boldsymbol{y}^s - m, \quad s \in \mathcal{S} \\ & \boldsymbol{w} \in W. \end{array}\right. \tag{4.14}$$

According to [74], the β-CVaR approximates the conditional expectation of the loss exceeding the β-VaR, and for β close to one, the solution to Problem (4.14) is expected to be similar to that to Problem (4.13).

All computations were conducted on a personal computer with Pentium4 processor (3.4 GHz) and 2 GB memory. MATLAB R2006b with optimization toolbox was used to implement the proposed algorithms and the convex approximation, while Xpress-MP release 2006B was used for the LP (4.14) for the CVaR minimization and the MIP formulation.

Tables 4.1 (i) to (iii) show the computational results for five-asset problems. Each table corresponds to different number of scenarios S=100, 1,000 and 10,000. All the figures show the average of five experiments, each using different scenario set, but generated from the identical distribution mentioned above. When S=100, the MIP formulation quickly achieves the VaR in an exact manner. However, when S=1,000 and 10,000, the MIP formulation cannot be solved within 10 hours or results in memory shortage. On the other hand, the proposed algorithms attain better solutions than those of the convex approximation (c) and the CVaR minimization (d). Moreover, if approximation accuracy α is relaxed from 2 to 5, CPU time decreases sharply whereas the difference of the achieved VaRs is small. CPU time of the proposed algorithms does not increase so much even when S grows from 1,000 to 10,000. It may be worth mentioning that when approximation accuracy α is small or the number of assets is larger, the Kelley's method (b) is expected to be superior to the linear relaxation (a). This is because the size of branch-and-bound tree becomes larger owing to the excessively relaxed linear relaxation. It is also observed that the resulting VaR of the convex approximation is no less than that of the CVaR minimization in all the results.

Figure 4.2 plots how VaRs for the incumbent solutions are updated, where the proposed algorithm with the Kelley's cutting plane method and α=2, and the MIP formulation are compared through five experiments with S = 10,000 and N = 5. The vertical axis shows VaR($\bar{\boldsymbol{w}}^k_{\mathrm{VaR}}$) for the VaR best solution $(\bar{\boldsymbol{w}}^k_{\mathrm{VaR}}, \bar{m}^k_{\mathrm{VaR}})$ of the proposed algorithm and VaR for the incumbent solution of the MIP formulation. In all tests, incumbent VaRs of the proposed algorithm fall sharply at an early stage, and therefore, we see that the proposed algorithm works efficiently as a heuristics by quitting computation before satisfying optimality. For example, in the scenario set No.2, although the proposed algorithm spent 16,000 seconds to meet the optimality,

Table 4.1: The VaR, the Violation Probability, and the Computation Time ($N = 5$)

The column "VaR" displays the value of VaR($\boldsymbol{w}^*$) for the obtained solution ($\boldsymbol{w}^*, m^*$) via each approach, while the column "VP" displays the violation probability VP($\boldsymbol{w}^*, m^*$) := Prob$\{(\boldsymbol{w}^*)^\top \tilde{\boldsymbol{y}} - m^* > 0\}$.

(i) S=100

	VaR	VP	CPU time (sec)
(a) SBB with linear relaxation, $\alpha = 2$	3.53	0.076	408.1
(a) SBB with linear relaxation, $\alpha = 5$	3.57	0.044	57.3
(b) SBB with Kelley's method, $\alpha = 2$	3.45	0.074	573.1
(b) SBB with Kelley's method, $\alpha = 5$	3.58	0.040	80.8
(c) Convex Approximation	5.18	0.038	0.3
(d) CVaR minimization	4.89	-	0.1
(e) MIP formulation	3.24	0.100	2.1

(ii) S=1,000

	VaR	VP	CPU time (sec)
(a) SBB with linear relaxation, $\alpha = 2$	4.04	0.070	6816.3
(a) SBB with linear relaxation, $\alpha = 5$	4.08	0.050	239.4
(b) SBB with Kelley's method, $\alpha = 2$	4.03	0.071	6332.6
(b) SBB with Kelley's method, $\alpha = 5$	4.10	0.047	386.9
(c) Convex Approximation	5.26	0.037	0.3
(d) CVaR minimization	5.22	-	0.1
(e) MIP formulation	-	-	over 10 hours

(iii) S=10,000

	VaR	VP	CPU time (sec)
(a) SBB with linear relaxation, $\alpha = 2$	4.38	0.073	4730.9
(a) SBB with linear relaxation, $\alpha = 5$	4.40	0.048	331.7
(b) SBB with Kelley's method, $\alpha = 2$	4.38	0.073	6539.1
(b) SBB with Kelley's method, $\alpha = 5$	4.39	0.050	618.5
(c) Convex Approximation	5.45	0.039	0.5
(d) CVaR minimization	5.45	-	0.9
(e) MIP formulation	-	-	memory shortage

the optimal solution was obtained within 2,000 seconds. Furthermore, an incumbent solution which has the nearly optimal objective value is obtained within only 10 seconds. MIP formulation results in memory shortage, and incumbent VaRs are much larger than those of the proposed algorithm.

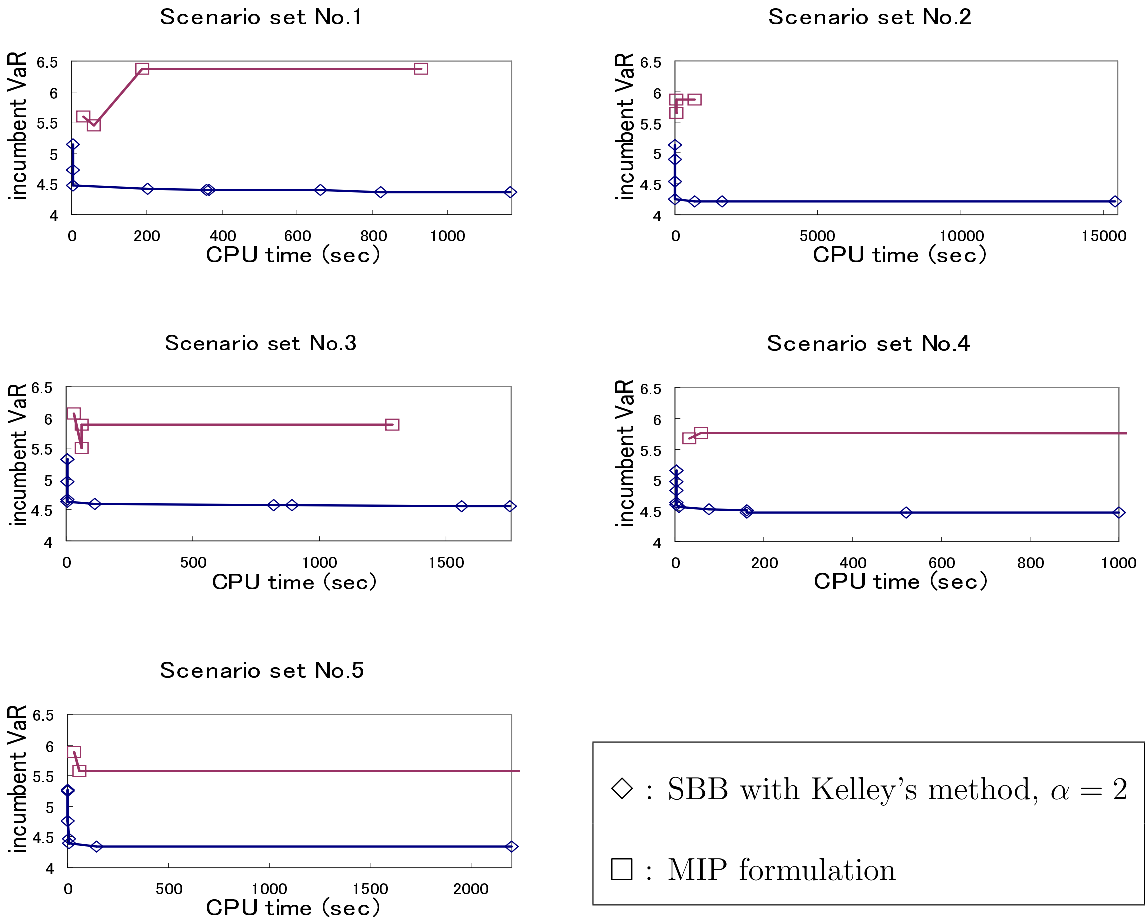


Figure 4.2: Updated Value-at-Risk for Incumbent Solutions with respect to CPU Time

Table 4.2 shows the results of $N = 10$ and $S = 10{,}000$. The proposed algorithm uses Kelley's cutting plane method for computing a lower bound on a simplex. We set $\alpha = 10$. In Table 4.2, the proposed algorithm is used as a heuristics by stopping the computation in 100 seconds. As a result, the resulting VaR of the proposed algorithm is rather smaller than that of the other approximation methods in all the tests.

Figures 4.3 (a) to (c) show the portfolios $\boldsymbol{w}^*$ obtained via the α-conservative

Table 4.2: The VaR and the Violation Probability via the Three Approaches ($N = 10$ and $S = 10{,}000$)

The columns corresponding to "SBB ($\alpha = 10$)" shows the results of the proposed branch-and-bound algorithm with $\alpha = 10$ where Kelley's cutting plane method is used for solving convex subproblems and the algorithm terminates in 100 seconds. The columns corresponding to "Convex App." shows the results via the convex conservative approach (c), while "CVaR min." shows those via the CVaR minimization approach (d).

	SBB ($\alpha = 10$)		Convex App.		CVaR min.	
Scenario set	VaR	VP	VaR	VP	VaR	VP
No. 1	3.86	0.027	3.95	0.039	3.95	-
No. 2	3.68	0.028	4.02	0.040	4.02	-
No. 3	3.79	0.029	4.03	0.041	4.03	-
No. 4	3.65	0.020	4.04	0.041	4.03	-
No. 5	3.73	0.029	4.13	0.041	4.12	-
Average	3.74	0.027	4.03	0.040	4.03	-

approach with $\alpha = 2$ for the five-scenario sets. From these figures, we see that the solutions for 100 scenarios are so scattered that they are not reliable, while the solution for 10,000 scenarios are rather stable. This fact explains the motivation for increasing the scenario size S so as to ensure the reliability of the obtained portfolio.

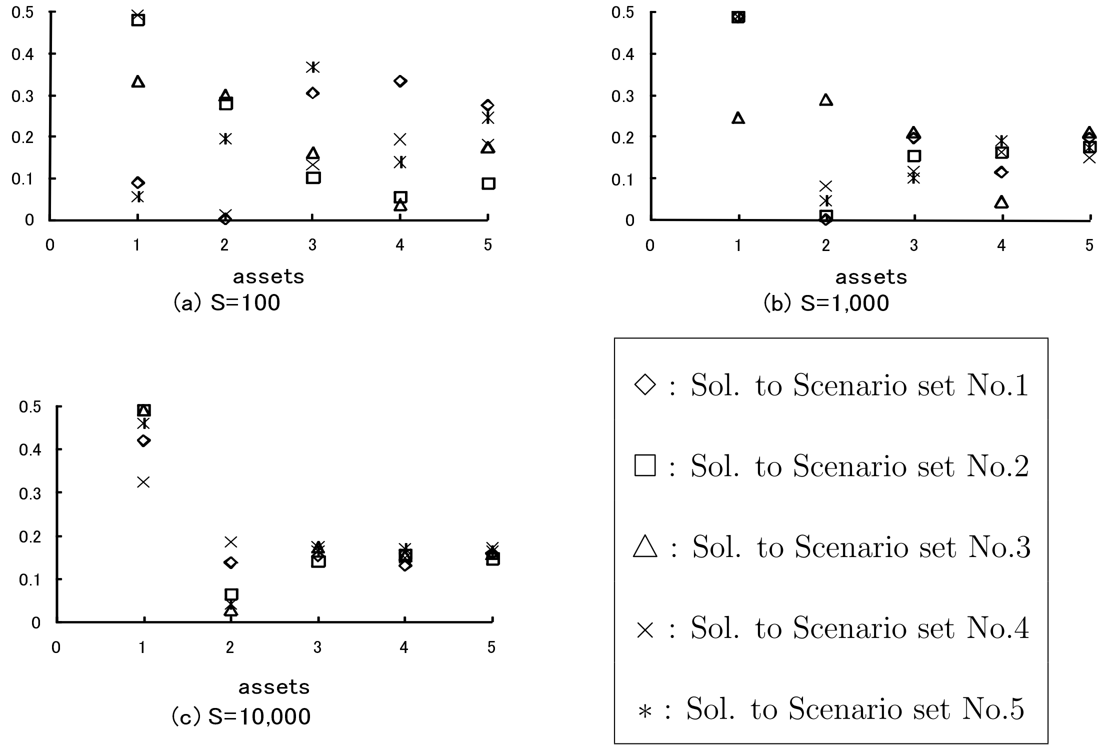


Figure 4.3: Optimal investment ratio to five-scenario sets when five assets are considered

Chapter 5

Constant Rebalanced Portfolio Optimization under Nonlinear Transaction Costs

5.1 Constant Rebalancing under Transaction Costs

We start with giving a mathematical description of a constant rebalancing model under transaction costs.

Let us define the index sets as follows:

$\mathcal{N} := \{1, ..., N\}$	:	index set of investable financial assets
$\mathcal{T} := \{1, ..., T\}$	:	index set of time periods in the future
$\mathcal{S} := \{1, ..., S\}$	:	index set of given scenarios, or in other words, index set of simulated paths (see Figure 5.2).

Let $\gamma_i : \mathbb{R} \to \mathbb{R}$ be a function representing the transaction cost of asset $i \in \mathcal{N}$ for the amount of transaction η. We assume that γ_i is the following convex function representing a market impact cost:

$$\gamma_i(\eta) := a_i \,[\,-\eta\,]^+ \exp\left(-b_i\eta\right) + c_i \,[\,\eta\,]^+ \exp\left(d_i\eta\right), \tag{5.1}$$

where a_i, b_i, c_i and d_i are non-negative parameters to be estimated. The value of γ_i is almost zero in case of small amount of transaction and grows exponentially as the amount of transaction increases. The widely used linear transaction cost e.g., [11, 18, 70, 78] can be represented by (5.1) with $b_i = d_i = 0$.

Figure 5.1 illustrates an example of constant rebalancing. Let us assume that 4.0, 3.0 and 3.0 billion Japanese yen are invested in Stock A, Stock B and Bond C respectively at the beginning of the planning horizon. At the end of the first period, investment proportion will differ from the initial proportion due to the change of asset prices. Then, the constant rebalancing strategy compels to purchase assets whose price has decreased and to sell assets whose price has increased so that the proportion is restored to be 40%, 30% and 30% at the beginning of the next period.

Let

$\bar{y}_i$: initial investment unit in asset i (given parameters, $i \in \mathcal{N}$)

$p_{0,i}$: initial price of asset i per unit (given parameters, $i \in \mathcal{N}$)

$p^s_{t,i}$: price of asset i per unit at the end of period t under the scenario s (given parameters, $t \in \mathcal{T},\ s \in \mathcal{S},\ i \in \mathcal{N}$)

$d^s_{t,i}$: dividend of asset i per unit held at the end of period t under the scenario s (given parameters, $t \in \mathcal{T},\ s \in \mathcal{S},\ i \in \mathcal{N}$)

v_0 : initial wealth (given parameter)

v^s_t : portfolio value before rebalancing at the end of period t under the scenario s (decision variables, $t \in \mathcal{T},\ s \in \mathcal{S}$)

u_0 : portfolio value after the initial rebalancing (decision variable)

u^s_t : portfolio value after rebalancing at the end of period t under the scenario s (decision variables, $t \in \mathcal{T} \setminus \{T\},\ s \in \mathcal{S}$)

w_i : investment proportion in asset i (decision variables, $i \in \mathcal{N}$).

We assume that the investor has an initial portfolio $\bar{\boldsymbol{y}}$ at the beginning of the investment. The constant rebalancing strategy enforces the rebalancing to the proportion $\boldsymbol{w}$ at the beginning of each discrete investment period. In case of initial rebalancing, the invested amount $p_{0,i}\bar{y}_i$ is adjusted to $u_0 w_i$, and at the same time, portfolio value u_0 is calculated by subtracting the transaction cost from initial wealth

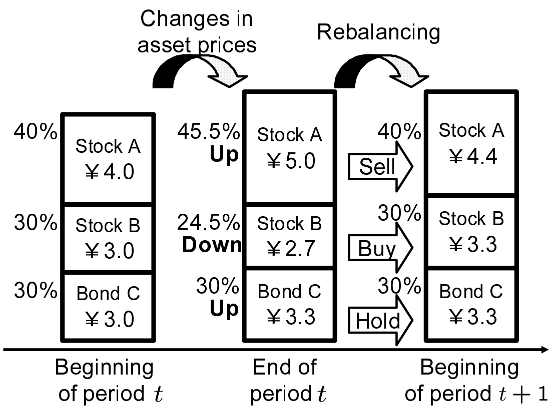


Figure 5.1: Example of Constant Rebalancing

v_0. The relation between initial wealth v_0 and portfolio value u_0 is given by

$$v_0 = u_0 + \sum_{i \in \mathcal{N}} \gamma_i (u_0 w_i - p_{0,i} \bar{y}_i). \tag{5.2}$$

Due to the asset price changes and receipt of dividends, the portfolio value changes over the first period. Accordingly, portfolio value before rebalancing at the end of the first period under the scenario s is given by

$$v_1^s = u_0 \sum_{i \in \mathcal{N}} (1 + r_{1,i}^s) w_i, \tag{5.3}$$

where $r_{t,i}^s$ is

$$r_{1,i}^s := \frac{p_{1,i}^s - p_{0,i} + d_{1,i}^s}{p_{0,i}}, \tag{5.4}$$

and represents the return of asset i at the period t under the scenario s. In general, $r_{t,i}^s$ is given by

$$r_{t,i}^s := \frac{p_{t,i}^s - p_{t-1,i}^s + d_{t,i}^s}{p_{t-1,i}^s}, \quad t \in \mathcal{T} \setminus \{1\}. \tag{5.5}$$

Taking it into consideration that the investment unit in asset i right before the rebalancing is $(u_0 w_i)/p_{0,i}$, in the same way as Equation (5.2), the relation between portfolio values v_1^s and u_1^s under the scenario s is given by

$$v_1^s = u_1^s + \sum_{i \in \mathcal{N}} \gamma_i \left(u_1^s w_i - \frac{p_{1,i}^s u_0 w_i}{p_{0,i}} \right). \tag{5.6}$$

Similarly, portfolio value before rebalancing at the end of period $t \in \mathcal{T} \setminus \{1\}$ under the scenario s is given by

$$v_t^s = u_{t-1}^s \sum_{i \in \mathcal{N}} (1 + r_{t,i}^s) w_i, \tag{5.7}$$

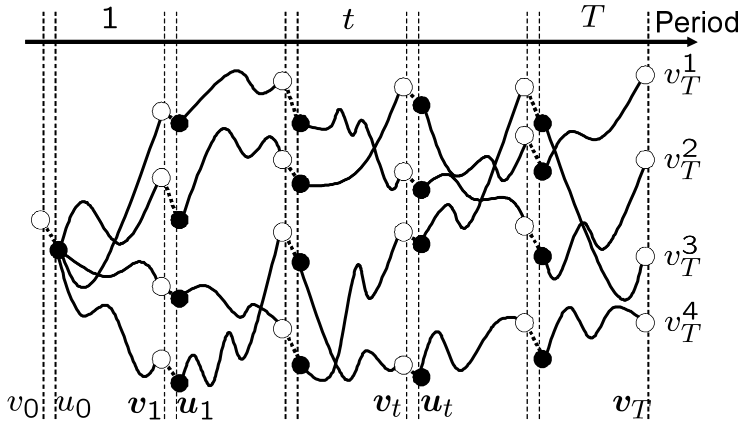


Figure 5.2: Simulated Paths of Portfolio Value

and the relation between portfolio values v_t^s and u_t^s at the end of period $t \in \mathcal{T} \setminus \{1, T\}$ under the scenario s is given by

$$v_t^s = u_t^s + \sum_{i \in \mathcal{N}} \gamma_i \left(u_t^s w_i - \frac{p_{t,i}^s u_{t-1}^s w_i}{p_{t-1,i}^s} \right). \tag{5.8}$$

In Figure 5.2, an example of changes in portfolio value is illustrated. The portfolio value falls at the beginning of each period due to transaction costs associated with the rebalancing (see Equations (5.2), (5.6) and (5.8)).

5.2 Portfolio Optimization Problem

In this section, we formulate a constant rebalanced portfolio optimization problem in which expected return is maximized subject to CVaR constraints.

5.2.1 Formulation

In the rest of this section, we assume that occurrence probability of scenario $s \in \mathcal{S} = \{1, ..., S\}$ is $1/S$. We consider the following constraints:

Portfolio Dynamics Equations

$$\begin{cases} v_0 = u_0 + \sum_{i \in \mathcal{N}} \gamma_i (u_0 w_i - p_{0,i} \bar{y}_i), & & \cdots (5.9.\,\mathrm{a}) \\ v_1^s = u_1^s + \sum_{i \in \mathcal{N}} \gamma_i \left(u_1^s w_i - \dfrac{p_{1,i}^s u_0 w_i}{p_{0,i}} \right), & s \in \mathcal{S} & \cdots (5.9.\,\mathrm{b}) \\ v_t^s = u_t^s + \sum_{i \in \mathcal{N}} \gamma_i \left(u_t^s w_i - \dfrac{p_{t,i}^s u_{t-1}^s w_i}{p_{t-1,i}^s} \right), & t \in \mathcal{T} \setminus \{1, T\},\ s \in \mathcal{S} & \cdots (5.9.\,\mathrm{c}) \\ v_1^s = u_0 \sum_{i \in \mathcal{N}} (1 + r_{1,i}^s) w_i, & s \in \mathcal{S} & \cdots (5.9.\,\mathrm{d}) \\ v_t^s = u_{t-1}^s \sum_{i \in \mathcal{N}} (1 + r_{t,i}^s) w_i, & t \in \mathcal{T} \setminus \{1\},\ s \in \mathcal{S} & \cdots (5.9.\,\mathrm{e}) \end{cases} \tag{5.9}$$

Investment Proportion Constraints

$$\begin{cases} w_i^L \leq w_i \leq w_i^U, \quad i \in I & \cdots (5.10.\,\mathrm{a}) \\ \sum_{i \in \mathcal{N}} w_i = 1, & \cdots (5.10.\,\mathrm{b}) \end{cases} \tag{5.10}$$

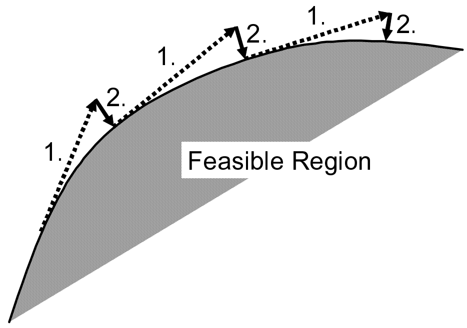
where w_i^L (w_i^U) is a lower (upper) limit of investment proportion in asset $i \in \mathcal{N}$.

In the following formulation, the random loss regarding the definition of CVaR is defined as "$(-1)\times$(portfolio value at the end of period T)," that is, $-v_T^s$, and both the maximization of the expected portfolio value at the end of period T and the minimization of CVaR are considered at the same time by taking the weighted sum of two objectives:

$$\left| \begin{array}{ll} \underset{\boldsymbol{u}, \boldsymbol{v}, \tau, \boldsymbol{w}}{\text{maximize}} & \lambda \left(\dfrac{1}{S} \sum_{s \in \mathcal{S}} v_T^s \right) - (1 - \lambda) \left(\tau + \dfrac{1}{(1 - \beta) S} \sum_{s \in \mathcal{S}} [-v_T^s - \tau]^+ \right) \\ \text{subject to} & \text{Portfolio Dynamics Equations (5.9)}, \\ & \text{Investment Proportion Constraints (5.10)}, \end{array} \right. \tag{5.11}$$

where $\lambda \in [0, 1]$ is the trade-off parameter between the expected return and the CVaR.

A large number of bilinear terms of decision variables ($u_0 w_i$, $u_t^s w_i$ and the like) appear in Constraint (5.9). Therefore, Problem (5.11) is nonconvex and difficult to



1. Solving a linear approximation problem, 2. Solving nonlinear equations

Figure 5.3: Geometric Interpretation of the Local Search Algorithm

attain a globally optimal solution in general [45, 81]. Moreover, when the number of scenarios is very large, the problem size becomes so huge that even a state-of-the-art NLP solver such as NUOPT may not provide any solutions. Thus in this thesis, we propose an iterative local search algorithm repeating the following two steps (Figure 5.3):

1. Solving a linear approximation problem for problem (5.11)

2. Solving nonlinear equations (5.9) via Newton's method

Fleten, Høyland and Wallace [32] stated that for their data sets, constant rebalanced portfolio optimization problems are virtually convex since their local search method using many different starting values for each instance always converged to the same solution. Although risk is measured by the expected accumulated quadratic shortfalls and transaction costs are not considered in their problem, their statement motivates us to apply a local search approach to Problem (5.11). Since we only solve an "approximation" problem for Problem (5.11), the obtained solution does not necessarily satisfy Constraint (5.9). Therefore, we need to find a feasible solution via Newton's method starting from a solution to the approximation problem.

In addition, our local search algorithm is suited to more general formulation such as the following problem with CVaR constraints. The random loss regarding the CVaR at each period is defined as "decrease in portfolio value during the period

t," that is, $v^s_{t-1} - v^s_t$, in the problem:

$$\left|\begin{array}{ll} \underset{\boldsymbol{u},\boldsymbol{v},\boldsymbol{\tau},\boldsymbol{w}}{\text{maximize}} & \dfrac{1}{S}\displaystyle\sum_{s\in\mathcal{S}} v^s_T \\ \text{subject to} & \tau_1 + \dfrac{1}{(1-\beta)S}\displaystyle\sum_{s\in\mathcal{S}} [v_0 - v^s_1 - \tau_1]^+ \leq \chi_1, \\ & \tau_t + \dfrac{1}{(1-\beta)S}\displaystyle\sum_{s\in\mathcal{S}} \left[v^s_{t-1} - v^s_t - \tau_t\right]^+ \leq \chi_t, \quad t \in \mathcal{T}\setminus\{1\} \\ & \text{Portfolio Dynamics Equations (5.9)}, \\ & \text{Investment Proportion Constraints (5.10)}, \end{array}\right. \tag{5.12}$$

where χ_t are user-defined parameters for representing upper bounds of CVaR value at the end of each period $t \in \mathcal{T}$. Considering portfolio value only at the end of the planning horizon is not enough since it is frequently uncertain when long-term investment will be discontinued. Imposing CVaR constraints on each rebalancing is, hence, meaningful in practice

5.2.2 Other Algorithms in the Literature

Maranas et al. [52] proposed a rectangular branch-and-bound algorithm for the constant rebalanced mean-variance portfolio optimization under no transaction costs. They enjoy the fact that their problem can be formulated as the simple one by eliminating the decision variables. (see Section 5.4.1 for detailed explanation). However, such a variable elimination is impossible under transaction costs, and therefore their algorithm cannot be applied to Problem (5.11). In addition, even if there are no transaction costs, the performance of the branch-and-bound procedure for Problem (5.12) deteriorates. This is because the number of variables and constraints of subproblems becomes large due to CVaR constraints.

A typical heuristic algorithm for the problem is an iterative optimization by alternately fixing decision variables, $\boldsymbol{w}$ and $\boldsymbol{u}$, which compose the bilinear terms. For instance, given $\bar{\boldsymbol{w}}$, Problem (5.11) is solved subject to $\boldsymbol{w} = \bar{\boldsymbol{w}}$, which is a convex program. Then, for the obtained solution $\bar{\boldsymbol{u}}$, Problem (5.11) is solved subject to $\boldsymbol{u} = \bar{\boldsymbol{u}}$, resulting in a new $\bar{\boldsymbol{w}}$. By repeating this procedure, the sequence of the obtained solutions is expected to improve. However, if decision variable $\boldsymbol{u}$ is fixed, degree of freedom of the other variables is excessively reduced as well. As a result, local search of Problem (5.11) subject to $\boldsymbol{u} = \bar{\boldsymbol{u}}$ provides a solution $\boldsymbol{w}$ which is almost the same as in the former iteration, and therefore the sequence of the obtained solutions is not improved enough. Hibiki [39] proposed an alternating optimization for approximately solving his model. His algorithm works well if a good initial solution is obtained. For the problem in hand, however, this alternating strategy fails to improve the incumbent solutions due to the excessive reduction of the feasibility of the subproblems. In addition, transaction costs are not considered in [39].

5.3 Local Search Algorithm

In this section, we explain a local search algorithm for solving Problem (5.11) in detail. The algorithm consists of two procedures: (i) solving a linear approximation problem, and (ii) finding a feasible solution.

5.3.1 Linear Approximation Problem for Problem (5.11)

The linear approximation problem for Problem (5.11), denoted by LAP($\bar{\boldsymbol{u}}$,$\bar{\boldsymbol{w}}$), is formulated as follows. First, the objective function is linearized by introducing auxiliary variables z_t^s. Next, the nonlinear terms of decision variables in Constraint (5.9) are linearly approximated at $(\bar{\boldsymbol{u}}, \bar{\boldsymbol{w}})$ with respect to $\boldsymbol{u}$, $\boldsymbol{w}$ (see (6.2) in Appendix C). As a result, LAP($\bar{\boldsymbol{u}}$,$\bar{\boldsymbol{w}}$) is the following problem:

$$
\text{LAP}(\bar{\boldsymbol{u}}, \bar{\boldsymbol{w}}) \left|
\begin{array}{lll}
\underset{\boldsymbol{u},\boldsymbol{v},\tau,\boldsymbol{w},\boldsymbol{z}}{\text{maximize}} & \lambda \left(\dfrac{1}{S} \displaystyle\sum_{s \in \mathcal{S}} v_t^s \right) - (1-\lambda) \left(\tau + \dfrac{1}{(1-\beta)S} \displaystyle\sum_{s \in \mathcal{S}} z_t^s \right) & \\
\text{subject to} & z_t^s \geq -v_t^s - \tau, \quad z_t^s \geq 0, & s \in \mathcal{S} \\
& \text{Linearly Approximated Portfolio Dynamics Equations (6.2)}, & \\
& \text{Investment Proportion Constraints (5.10)}, & \\
& \boldsymbol{w} \in \mathcal{W}(\bar{\boldsymbol{w}}), &
\end{array}
\right.
\tag{5.13}
$$

where $\mathcal{W}(\bar{\boldsymbol{w}})$ is a neighborhood of $\bar{\boldsymbol{w}}$, given by

$$\mathcal{W}(\bar{\boldsymbol{w}}) := \{\boldsymbol{w} \mid \bar{w}_i - \kappa_i \leq w_i \leq \bar{w}_i + \kappa_i,\ i \in \mathcal{N}\}, \tag{5.14}$$

where $\kappa_i > 0$ are step size parameters regarding investment proportion of asset $i \in \mathcal{N}$. The constraint $\boldsymbol{w} \in \mathcal{W}(\bar{\boldsymbol{w}})$ is imposed so that the solution to the linearly approximated problem will not be far from the incumbent point $(\bar{\boldsymbol{u}},\bar{\boldsymbol{w}})$.

5.3.2 Finding a Feasible Solution via Newton's Method

In order to linearly approximate Problem (5.11), we need a feasible $(\bar{\boldsymbol{u}},\bar{\boldsymbol{w}})$ at each iteration. However, a solution to LAP$(\bar{\boldsymbol{u}},\bar{\boldsymbol{w}})$ does not necessarily satisfy Constraint (5.9). In the following, we explain the procedure for finding a feasible $(\bar{\boldsymbol{u}},\bar{\boldsymbol{w}})$ by starting from a (possibly infeasible) solution of LAP$(\bar{\boldsymbol{u}},\bar{\boldsymbol{w}})$.

Let $(\boldsymbol{u}^{\mathrm{LAP}}, \boldsymbol{v}^{\mathrm{LAP}}, \tau^{\mathrm{LAP}}, \boldsymbol{w}^{\mathrm{LAP}}, \boldsymbol{z}^{\mathrm{LAP}})$ be a solution to LAP$(\bar{\boldsymbol{u}},\bar{\boldsymbol{w}})$. Obviously, $\boldsymbol{w}^{\mathrm{LAP}}$ is feasible to Problem (5.11) since Constraint (5.10) is included in LAP$(\bar{\boldsymbol{u}},\bar{\boldsymbol{w}})$. In the procedure explained below, we find a feasible solution $(\boldsymbol{w}^{\mathrm{LAP}}, \bar{\boldsymbol{u}}, \bar{\boldsymbol{v}})$ to Problem (5.11) by substituting $\boldsymbol{w}^{\mathrm{LAP}}$ into Constraint (5.9) and solving the nonlinear equations for $\boldsymbol{u}$ and $\boldsymbol{v}$.

1. Finding feasible u_0 via Constraint (5.9. a). In Constraint (5.9. a), if we substitute $\boldsymbol{w}^{\mathrm{LAP}}$, only u_0 is unknown. Considering that the right-hand side is convex in u_0, we can find $\bar{u}_0$ which satisfies Constraint (5.9. a) by applying Newton's method in u_0.

2. Finding feasible v_1^s $(s \in \mathcal{S})$ via Constraint (5.9. d). By substituting $\boldsymbol{w}^{\mathrm{LAP}}$ and $\bar{u}_0$ in Constraint (5.9. d), $\bar{v}_1^s$ $(s \in \mathcal{S})$ are determined.

3. Finding feasible u_1^s $(s \in \mathcal{S})$ via Constraint (5.9. b). Now in Constraint (5.9. b), only u_1^s $(s \in \mathcal{S})$ are unknowns. In the same manner as in case of Constraint (5.9. a), we find u_1^s $(s \in \mathcal{S})$ by applying Newton's method S times.

4. Finding feasible v_t^s $(t \in \mathcal{T} \setminus \{1\},\ s \in \mathcal{S})$ and u_t^s $(t \in \mathcal{T} \setminus \{1, T\},\ s \in \mathcal{S})$ via Constraints (5.9. e) and (5.9. c). In the same way as described above, by repeating the following procedures until $t = T$ is satisfied, we obtain a feasible solution $(\boldsymbol{w}^{\mathrm{LAP}},\bar{\boldsymbol{u}},\bar{\boldsymbol{v}})$ to Problem (5.11).

- By substituting $\boldsymbol{w}^{\mathrm{LAP}}$ and $\bar{u}^s_{t-1}$ $(s \in \mathcal{S})$ in Constraint (5.9. e), $\bar{v}^s_t$ $(s \in \mathcal{S})$ are determined.
- In Constraint (5.9. c), if we substitute $\boldsymbol{w}^{\mathrm{LAP}}$, $\bar{u}^s_{t-1}$ $(s \in \mathcal{S})$ and $\bar{v}^s_t$ $(s \in \mathcal{S})$, we find $\bar{u}^s_t$ $(s \in \mathcal{S})$ by applying Newton's method S times.

5. Evaluation of objective value. The objective value of Problem (5.11) is calculated from $\bar{v}^s_t$ $(s \in \mathcal{S})$ obtained in the above procedures.

5.3.3 Outline of the Local Search Algorithm

We are now in a position to describe the local search algorithm.

Local Search Algorithm for Problem (5.11)

Step 0. [Initialization.] Let $\bar{\boldsymbol{w}}$ be a feasible solution to Problem (5.11) and set the maximum number of iterations.

Step 1. [Newton's method.] Substitute $\bar{\boldsymbol{w}}$ in Equations (5.9), and find a feasible solution to Problem (5.11) by solving Equations (5.9) via Newton's method. Set the obtained solution as $(\bar{\boldsymbol{u}}, \bar{\boldsymbol{v}})$.

Step 2. [Termination check.] If the objective value is not improved or the maximum number of iterations is reached, terminate the algorithm with the best solution obtained so far. Otherwise, go to Step 3.

Step 3. [Linear approximation problem.] Solve the linear approximation problem LAP($\bar{\boldsymbol{u}}$,$\bar{\boldsymbol{w}}$), and let $(\boldsymbol{u}^{\mathrm{LAP}}, \boldsymbol{v}^{\mathrm{LAP}}, \tau^{\mathrm{LAP}}, \boldsymbol{w}^{\mathrm{LAP}}, \boldsymbol{z}^{\mathrm{LAP}})$ be a solution to it. Set $\bar{\boldsymbol{w}} \leftarrow \boldsymbol{w}^{\mathrm{LAP}}$, and go to Step 1.

5.4 Computational Results

In this section, we report computational results, evaluating the effectiveness of the proposed algorithm and the performance of the constant rebalancing strategy. All computations was conducted on a Windows XP personal computer with AMD Athlon 64 Processor (2.41GHz) and 2GB memory, and NUOPT (ver.10.1.4), a mathematical programming software package developed by Mathematical System, Inc., was used.

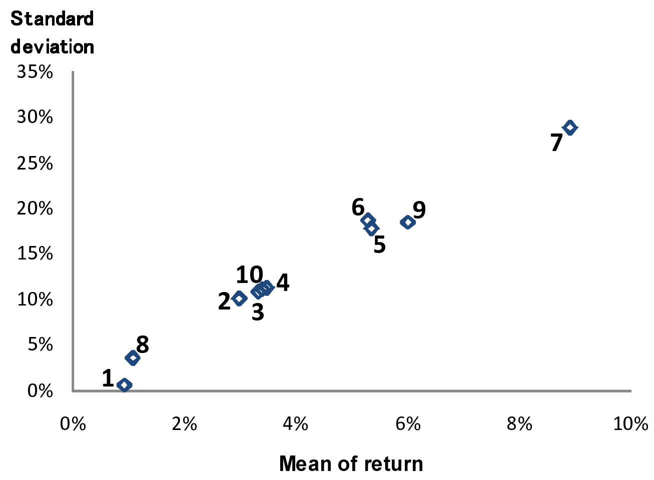


Figure 5.4: Mean and Standard Deviation of Asset Return

Problem Setting. Ten financial assets are considered over the planning horizon of five periods, and the number of scenarios (simulated paths) is 1,000, i.e., $N := 10$, $T := 5$ and $S := 1{,}000$. Asset $i = 1$ is cash with no transaction cost, and Assets $i = 2, 3, 4, 8, 10$ are low-risk assets with low transaction costs (e.g., bond), and Assets $i = 5, 6, 7, 9$ are high-risk assets with high transaction costs (e.g., stock) Figure 5.4 shows the Mean and Standard deviation of their returns. On the advice of Mizuho-DL Financial Technology Co., Ltd., the parameters of the transaction cost function γ_i are estimated using historical data, and the value of price and dividend in each scenario are generated via a bootstrap method. Lower limits of investment proportion are all zero (i.e., $\boldsymbol{w}^L := \boldsymbol{0}$), and the initial investment unit $\bar{\boldsymbol{y}}$ is set as $\bar{\boldsymbol{y}} := \boldsymbol{0}$, and the initial wealth v_0 is 1.0 trillion Japanese yen.

Parameter Setting of the Local Search Algorithm. Step size parameters κ_i are set as $\kappa_1 := 1$, $\kappa_i := 0.1$ $(i \in \mathcal{N} \setminus \{1\})$, and let the maximum number of iterations be eleven, i.e., the linear approximation problem is solved ten times.

5.4.1 Comparison with the Rectangular Branch-and-Bound Algorithm

Maranas et al. [52] proposed a rectangular branch-and-bound algorithm for globally solving the constant rebalanced mean-variance portfolio optimization. In this subsection, we revise their algorithm so that it solves mean-CVaR portfolio optimization, and compare the performance of the proposed local search algorithm with that of the revised version of the branch-and-bound algorithm.

Rectangular Branch-and-Bound Algorithm in [52]

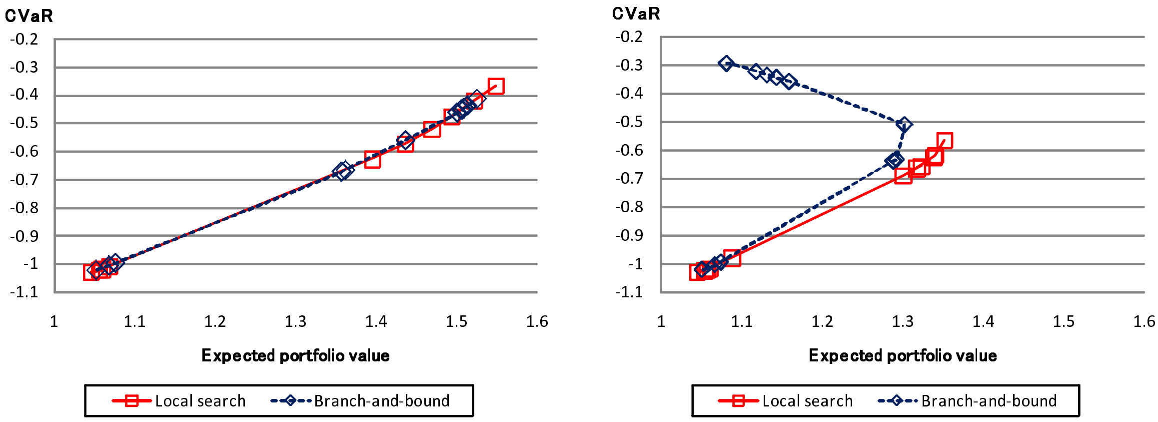
When no transaction cost is considered, one has $\boldsymbol{u} = \boldsymbol{v}$ in Constraints (5.9. a), (5.9. b) and (5.9. c), and accordingly, decision variable $\boldsymbol{u}$ can be eliminated. Then, by substituting $\boldsymbol{v}$ into the objective function, Problem (5.11) is reduced to the following formulation with $N+1$ variables and simple linear constraints:

$$\left|\begin{array}{ll} \underset{(\tau,\boldsymbol{w})\in\mathbb{R}\times\mathbb{R}^N}{\text{maximize}} & \lambda\left(\dfrac{1}{S}\displaystyle\sum_{s\in\mathcal{S}} v_0 \prod_{t\in\mathcal{T}}\left\{\sum_{i\in\mathcal{N}}(1+r^s_{t,i})w_i\right\}\right) \\ & -(1-\lambda)\left(\tau + \dfrac{1}{(1-\beta)S}\displaystyle\sum_{s\in\mathcal{S}}\phi\left(-v_0\prod_{t\in\mathcal{T}}\left\{\sum_{i\in\mathcal{N}}(1+r^s_{t,i})w_i\right\}-\tau\right)\right) \\ \text{subject to} & \text{Investment Proportion Constraints (5.10),} \end{array}\right. \tag{5.15}$$

where $\phi : \mathbb{R} \to \mathbb{R}$ is a smoothing function of nondifferentiable plus function $[\,\eta\,]^+$. In this thesis, we adopt

$$\phi(\eta) := \frac{\sqrt{\eta^2 + 4\delta^2} + \eta}{2}, \tag{5.16}$$

proposed in [66], where $\delta > 0$ is a parameter representing approximation accuracy. The rectangular branch-and-bound algorithm proposed in [52] is a solution method for the constant rebalanced mean-variance portfolio optimization under no transaction costs. See Appendix D for the revised version of convex subproblem over the subrectangle. The algorithm works well for problems of small number of assets N, but it starts deteriorating as N grows.



(a) Under no transaction costs (b) Under transaction costs

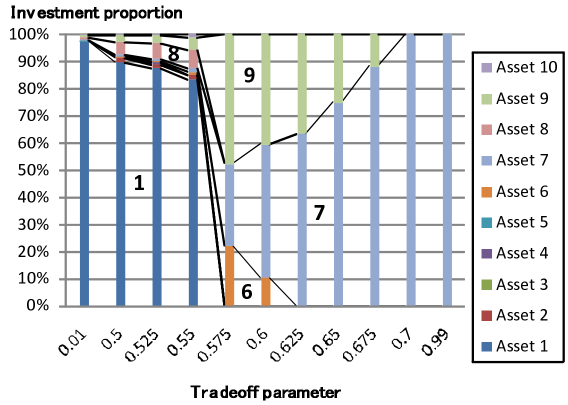
Figure 5.5: Efficient Frontier by Applying the Local Search Algorithm and the Rectangular Branch-and-Bound Algorithm

Discussion on Results

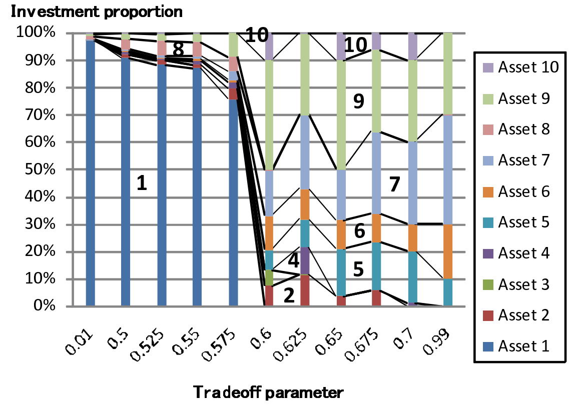
Efficient Frontier. Figure 5.5 shows the efficient frontier of the solutions obtained by the two algorithms. The horizontal axis is the expected portfolio value at the end of period T, that is $\frac{1}{S}\sum_{s\in\mathcal{S}} v_T^s$, and the vertical axis is the CVaR representing a risk of decrease in portfolio value at the end of period T, that is $\min\{\, \tau + \frac{1}{(1-\beta)S}\sum_{s\in\mathcal{S}}[-v_T^s - \tau\,]^+ \mid \tau \in \mathbb{R}\}$. Each plot corresponds to different value of λ chosen from $\{0.01, 0.5, 0.525, 0.55, 0.575, 0.6, 0.625, 0.65, 0.675, 0.7, 0.99\}$.

Figure 5.5 (a) depicts two kinds of efficient frontiers under no transaction costs. We see that the two kinds of frontiers almost coincide, which indicates that the local search algorithm attains almost optimal solutions since the branch-and-bound algorithm achieves the global optimality[1]. Figure 5.5 (b) shows the results under transaction costs. The efficient frontier of the branch-and-bound algorithm is drawn with the same solutions as those in Figure 5.5 (a), i.e., the solutions under no transaction costs. We see that the branch-and-bound algorithm can provide highly

[1]In Appendix D, the subproblems are proved to be convex when a parameter Θ is sufficiently large. It is, however, difficult to ascertain whether the parameter value is properly set (see [51] for the details), and accordingly, it is possible that a globally optimal solution may not be reached by the branch-and-bound algorithm in experimental results.



(a) Under no transaction costs



(b) Under transaction costs

Figure 5.6: Optimal Investment Proportion Provided by the Local Search Algorithm

inefficient solutions. This is due to the large transaction costs incurred. This result implies that neglecting transaction costs results in an insufficient investment in the presence of transaction costs.

Optimal Investment Proportion. Figure 5.6 shows the optimal investment proportion provided by the local search algorithm. Comparing the results under no transaction costs (Figure 5.6 (a)) with those under transaction costs (Figure 5.6 (b)), we observe that Asset 1 (cash) has the largest proportion when λ is small (i.e., low-risk investment) in both results. On the other hand, when λ is large (i.e., high-return investment), Asset 7 monopolizes the whole investments under no transaction costs, while investments are diversified among four or five assets under transaction costs.

5.4.2 Comparison with the Buy-and-Hold Strategy

Buy-and-hold strategy is a popular and simple investment strategy in which investor buys financial assets at the beginning and holds them until the end of planning horizon without trading. In this subsection, we compare the performance of this strategy with that of the constant rebalancing strategy, both of which take the transaction costs into account.

Buy-and-Hold Strategy

Portfolio optimization problem with buy-and-hold strategy under transaction costs is formulated as the following problem:

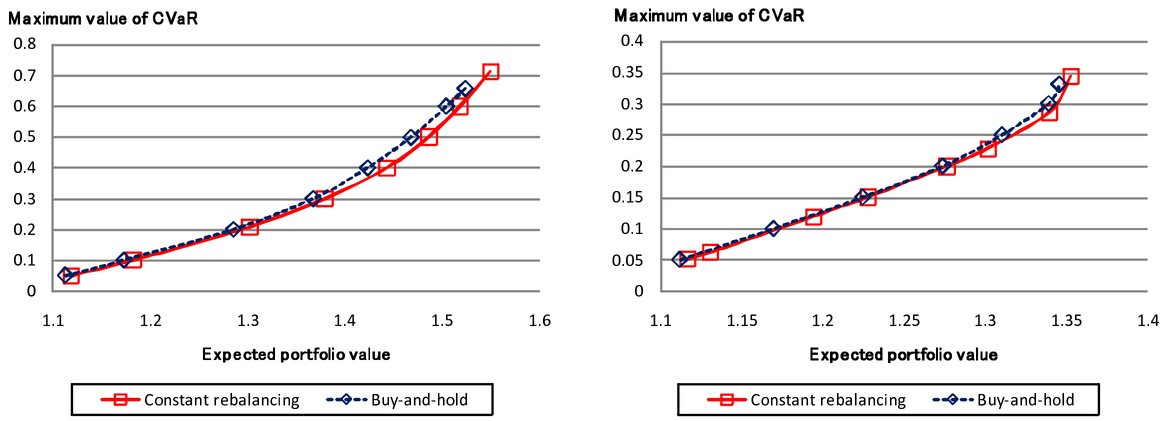
$$
\left|
\begin{array}{lll}
\underset{\boldsymbol{\tau},\boldsymbol{v},\boldsymbol{y}}{\text{maximize}} & \dfrac{1}{S}\displaystyle\sum_{s\in\mathcal{S}} v_T^s & \\
\text{subject to} & \tau_1 + \dfrac{1}{(1-\beta)S}\displaystyle\sum_{s\in\mathcal{S}} [v_0 - v_1^s - \tau_1]^+ \leq \chi_1, & \\
& \tau_t + \dfrac{1}{(1-\beta)S}\displaystyle\sum_{s\in\mathcal{S}} \left[v_{t-1}^s - v_t^s - \tau_t\right]^+ \leq \chi_t, & t \in \mathcal{T}\setminus\{1\} \\
& \displaystyle\sum_{i\in\mathcal{N}} p_{0,i}\, y_i + \sum_{i\in\mathcal{N}} \gamma_i(p_{0,i}y_i - p_{0,i}\bar{y}_i) \leq v_0, & \\
& v_t^s = \displaystyle\sum_{i\in\mathcal{N}} \left(p_{t,i}^s y_i + \sum_{\theta=1}^{t} d_{\theta,i}^s y_i\right), & t\in\mathcal{T},\ s\in\mathcal{S} \\
& y_i^L \leq y_i \leq y_i^U, & i\in\mathcal{N},
\end{array}
\right. \tag{5.17}
$$

where y_i denotes investment unit in asset $i \in \mathcal{N}$, and y_i^L (y_i^U) is a lower (upper) limit of investment unit in asset $i \in \mathcal{N}$. In this subsection, lower limits of investment unit are all zero (i.e., $\boldsymbol{y}^L := \mathbf{0}$). Since the transaction cost γ_i is the convex function (5.1), we apply NUOPT to Problem (5.17). In this problem, the decision variables for representing portfolio are not investment proportion $\boldsymbol{w}$, but investment unit $\boldsymbol{y}$, and portfolio value v_t^s ($t \in \mathcal{T},\ s \in \mathcal{S}$) is defined as the sum of asset values in market at the end of period t and dividends obtained by the end of period t.

Discussion on Results

Efficient Frontier. Figure 5.7 shows the efficient frontier of the two strategies. Upper bounds χ_t of CVaR are fixed to χ for all $t \in \mathcal{T}$, and we choose χ from {0.05, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 1} in case of no transaction costs and from {0.05, 0.1, 0.15, 0.2, 0.25, 0.3, 0.35, 1} in case of transaction costs. Each plot in Figure 5.7 corresponds to the solutions of Problem (5.12) via our local search algorithm[2] and Problem (5.17) via the NLP solver, respectively. Then, the horizontal axis is the

[2] Although the proposed algorithm does not necessarily provide a solution satisfying the CVaR constraints, the provided solution in this computational results almost satisfied them.



(a) Under no transaction costs (b) Under transaction costs

Figure 5.7: Efficient Frontier of the Constant Rebalancing Strategy and the Buy-and-Hold Strategy

expected portfolio value at the end of period T, and the vertical axis is the maximum value of CVaR with respect to period $t \in \mathcal{T}$.

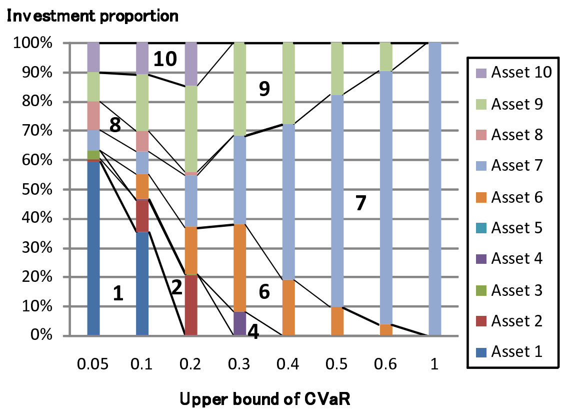
Comparing the results under no transaction costs (Figure 5.7 (a)), the two kinds of frontiers are almost the same when low-risk investment is made. On the other hand when high-return investment is made, the constant rebalancing strategy dominates the buy-and-hold strategy. One reason for this is that the obtained dividends can be invested in high-return assets in case of constant rebalancing. When transaction costs are incurred (Figure 5.7 (b)), the frontier provided by the constant rebalancing strategy dominates that by the buy-and-hold strategy on the right half of the figure (i.e., high-return investment); however, the difference is smaller than that under no transaction costs. This is because of the transaction costs that the investor should pay to rebalance the portfolio.

Optimal Investment Proportion. Figure 5.8 shows the optimal investment proportion of the constant rebalancing strategy and the buy-and-hold strategy. The two strategies provide similar investment proportion under no transaction costs (Figure 5.8 (a) and (c)) and also under transaction costs (Figure 5.8 (b) and (d)). The smaller the upper bound of CVaR is, the larger the investment proportion in cash becomes, and the larger the upper bound of CVaR is, the larger the investment proportions in high-risk assets become. In addition, whereas investment proportion in Asset 7 is very large under no transaction costs, investments are diversified among four to six assets under transaction costs.

Computational Time. To draw the efficient frontier in Figure 5.7 (a) or (b), eight problems are solved with different parameter values. For drawing the efficient frontier of the constant rebalancing strategy, we sequentially solve the problems by gradually increasing the parameter of the upper bound of CVaR, χ. The obtained solution under the previous χ is employed as the initial solution to the problem for the next χ. First, starting value of investment proportion, $\bar{\boldsymbol{w}}$, is set as $\bar{w}_1 := 1,\ \bar{w}_i := 0\ (i \in \mathcal{N}\backslash\{1\})$, and the lowest risk investment problem (say, $\chi = 0.05$) is solved. This investment proportion $\bar{\boldsymbol{w}}$ satisfies Constraint (5.10) and probably CVaR constraints of Problem (5.12) since Asset 1 is cash which is the lowest risk asset. Next, the obtained solution is employed as a starting value, $\bar{\boldsymbol{w}}$, and the second lowest risk investment problem (say, $\chi = 0.1$) is solved. Repeating this procedure to the highest return investment problem (say, $\chi = 1$), an efficient frontier is drawn with the small number of iterations of the algorithm. In the experiments, the algorithm for constant rebalancing terminates with three iterations, i.e., the linear approximation problem is solved only twice for all the problems under transaction costs; the average CPU time is 287.6 seconds, whereas the average CPU time of the buy-and-hold strategy is 18.5 seconds.

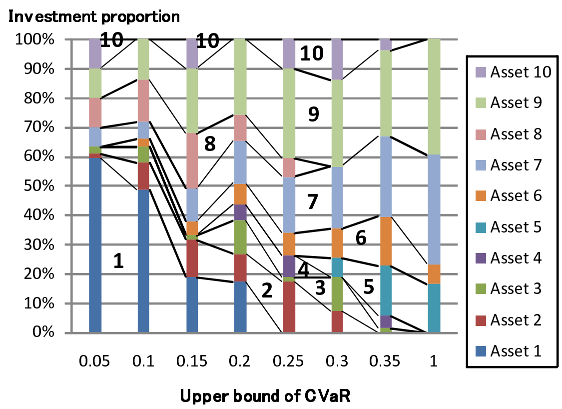
5.4.3 Out-of-Sample Performance

In this subsection, we conduct experiments for evaluating the out-of-sample performance of the constant rebalancing strategy under transaction costs. Scenario sets A and B, each containing 1,000 scenarios, are generated via a bootstrap method using the same historical data. Figure 5.9 shows two kinds of efficient frontiers where the setting is the same as that in Figure 5.7 (b). In the results using the scenario set B (Figure 5.9 (b)), the frontier of out-of-sample solutions differs from that of in-sample solutions at high-return points, however the two frontiers are almost the same. Then, although it has been shown in recent papers (e.g., [26]) that equally weighted portfolio performs well in out-of-sample tests, the performance of equally weighted portfolio is dominated by that of the optimal proportions. This is common in the observation reported in [32] that the constant rebalancing approach performs better in the out-of-sample result than in the in-sample result compared to the stochastic dynamic approach.



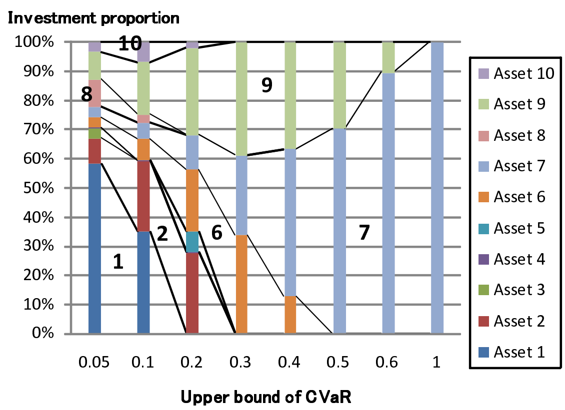
(a) Constant rebalancing

(under no transaction costs)

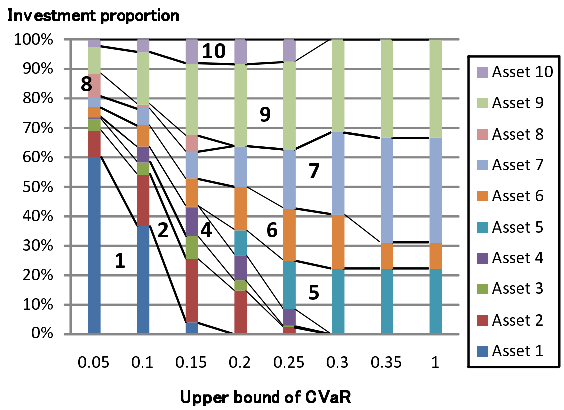


(b) Constant rebalancing

(under transaction costs)

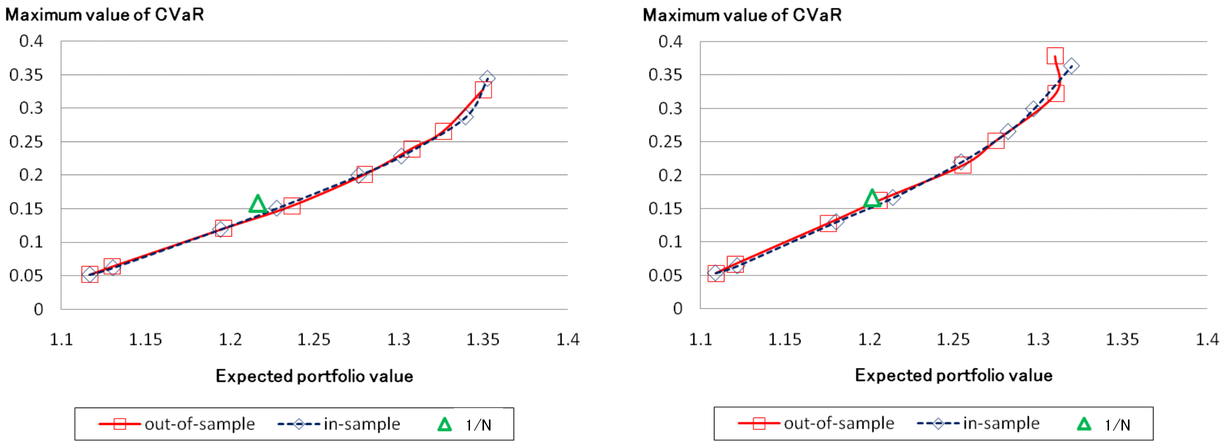


(c) Buy-and-hold (under no transaction costs)



(d) Buy-and-hold (under transaction costs)

Figure 5.8: Optimal Investment Proportion of the Constant Rebalancing Strategy and the Buy-and-Hold Strategy



(a) Results of scenario set A (b) Results of scenario set B

[out-of-sample] an optimal investment proportion of Problem (5.12) using the scenario set which is different from what we use for evaluating the performance; [in-sample] an optimal investment proportions of Problem (5.12) using the same scenario set with what we use for evaluating the performance; [$1/N$] an equally weighted portfolio, that is $\boldsymbol{w} = (1/N, 1/N, ..., 1/N)$.

Figure 5.9: Out-of-Sample Performance of the Constant Rebalancing Strategy under Transaction Costs

Chapter 6

Concluding Remarks

In this thesis, we have considered three types of stochastic programming problems: "Newsvendor Solutions via Conditional Value-at-Risk Minimization", "α-Conservative Approximation for Probabilistically Constrained Convex Programs" and "Constant Rebalanced Portfolio Optimization under Nonlinear Transaction Costs". Under uncertainty, we have to take into consideration the risk of incurring a large loss. One of the our contributions has been to build risk averse models via risk measure VaR/CVaR for the typical stochastic programming problems i.e., newsvendor problem and (single/multi-period) portfolio optimization problem. Moreover, we have proposed solution methods for these problems. It has been found from the experimental results that our methods work efficiently at some settings, however we need to improve the performance of solution algorithms for utilizing these models in a practical situation. Further research would clarify the effectiveness of stochastic programming approach to decision making problems under uncertain environments.

6.1 Newsvendor Solutions via Conditional Value-at-Risk Minimization

In this thesis, we apply the two different CVaR measures to the classic single-period newsvendor problem by introducing two loss functions called as the net loss and the total cost. Each of these measures captures a risk of the profit going down or the cost going up, respectively, to a certain level in a predetermined significance, and both are found to be convex functions with respect to the order quantities of products.

We demonstrate in this thesis that their convex structure plays an important role in seeking optimal solutions of the problems which contain the CVaR measure in objective and constraints as in the mean-CVaR models. Actually, one can achieve closed form solutions of the two CVaR minimizations and some mean-CVaR model in single product case where no constraint is imposed. Through some analysis, we see that the two CVaR minimizers show some differences in the order quantities and the solution sensitivity to associated parameters when $\beta > 0$, while both of them provide the same solution as the classic expected profit maximizer when $\beta = 0$. More specifically, the net loss CVaR minimizer leads to less order quantity than the classic solution when a shortage penalty parameter is set to be sufficiently small.

Even in the case of multiple products under linear constraints, one can compute a solution by solving a linear program if their demand distribution is given by a finite number of scenarios. This scenario based approach can be exploited for approximating an optimal solution under a certain distribution. Analysis of its statistical properties including the confidence intervals of the obtained values is yet to be explored. Furthermore, preliminary computational experiments show the efficiency of the LP solutions.

As Khouja [42] reviewed, the newsvendor problem has many directions for extension. Instead of the profit maximization, wide range of applications of the CVaR minimization to such advanced settings where risk attitude should be incorporated are expected.

6.2 α-Conservative Approximation for Probabilistically Constrained Convex Programs

In this thesis, we construct the α-conservative approximation problem of the probabilistically constrained convex program (PCCP), and show that it can be formulated as a D.C optimization problem. It is advantageous that the number of (sampled) scenarios does not affect the number of variables or constraints while it does not hold in the MIP formulation which requires a number of 0-1 variables, each corresponding to one scenario. Recalling the fact that more than ten thousands scenarios are required for sufficient accuracy of a solution even when the dimension of the solution is five, this property of the D.C. formulation is very preferable.

Although solving a genuine D.C. problem in a deterministic manner is known to be very hard, several algorithms are shown to have a potential to solve the problem especially when the number of variables concerned with the nonconvexity is up to ten (see, e.g., [45]).

In this thesis, the simplicial branch-and-bound method which is a famous deterministic algorithms for achieving a globally optimal solution is mainly investigated, and is applied to the VaR minimization of a financial portfolio.

Through the numerical experiments, we show that

- When the number of investable assets is up to five, a good (approximate) solution can be achieved in a practical amount of time for all cases by setting the parameter α to be two or five. It contrasts with the fact that the MIP formulation cannot solve the problem within ten hours or results in memory shortage when the number of scenarios is more than thousand. In addition, the convergence of the upper bound is much slower than the presented branch-and-bound algorithm for the D.C. optimization.

- By quiting computation before satisfying optimality, the proposed algorithm can be used as a heuristics. Although the solution quality deteriorates in terms of the original problem because of the abort, the resulting VaR is still smaller than that of the convex approximation and the CVaR optimization when the number of assets is ten.

- Convex approximation approach proposed by Nemirovski and Shapiro [63] provides a bit too conservative solution in terms of the VaR minimization (i.e., small violation probability), and the VaR at optimality is larger than that of CVaR minimization. Comparing to those convex approaches, the proposed nonconvex approach improves the quality of the solution in a practical time when the number of assets is five to ten.

Although, when the number of assets is hundred or more, the problem clearly becomes (prohibitively) hard and this nonconvex approach may not look appealing, it is worth noting that the number of scenarios is critical for the accuracy of the solution, and the number of scenarios required for sufficient accuracy drastically increases as the number of assets grows. Furthermore, in a practical situation which institutional investors face, such a portfolio selection is usually carried out in two (or more) steps, which is known as the asset allocation strategy. Typically, this strategy first allocates the fund into several asset classes, e.g., domestic stock, domestic bond, international stock, real estate, cash and so on. After the first allocation, the money allocated to each class is further allocated into individual assets within the class. In such a practical case, the number of classes or assets in each class is usually small and the proposed approach is expected to play a role in minimizing the VaR, ensuring high reliability.

6.3 Constant Rebalanced Portfolio Optimization under Nonlinear Transaction Costs

In this thesis, we formulate the constant rebalanced portfolio optimization problem under nonlinear transaction costs, and propose a solution method based on a local search approach. This problem is a nonconvex optimization including a large number of bilinear terms of decision variables in a number of constraints and difficult to attain a globally optimal solution in general. When a huge number of scenarios are considered, it becomes further difficult to attain a locally optimal solution via a state-of-the-art NLP solver. Then, we propose an iterative local search algorithm based on LP solution, which is easily attained even if the problem size is large, and on Newton's method for solving nonlinear equations. In the computational results, the proposed local search algorithm attains as good solution as the global optimization approach. Then, we show that the problem under transaction costs needs to be solved so as to obtain an efficient solution. Moreover, we see that the constant rebalancing strategy outperforms the buy-and-hold strategy when high return is sought. Furthermore in the out-of-sample performance, the constant rebalancing strategy is superior to the equally weighted portfolio. In addition, it should be noted that the proposed local search algorithm can deal with general nonlinear transaction costs.

Future tasks include improving the efficiency of the solution method and comparing the out-of-sample performance of the constant rebalancing strategy with that of the other various strategies. In a practical situation, the problem with a large number of scenarios (e.g., 10,000 or 100,000 scenarios) is desired to be solved. Therefore, we need to improve the algorithm for solving the problem with a large number of scenarios. On the other hand, dynamic stochastic approach e.g., [40], hybrid model [38, 39] and the like are not tested for comparing with the constant rebalancing strategy in this thesis. Although these are partly conducted in [32, 39], more detailed and inclusive comparison is essential for showing the effectiveness of the constant rebalancing strategy especially in out-of-sample performance.

Appendix

A Proof of Proposition 3.4

From (3.6) and (3.7), we note that the net loss $-\mathcal{P}(x,\xi)$ is represented as $V\,\xi - E\,(\,x-\xi\,)$ for $\xi \in [0,x]$, and $V\,\xi - U\,(\,\xi - x\,)$ for $\xi \in [x,\infty)$. Thus, the integral part of the objective in (3.8) can be expanded as

$$\int_0^x [-\{\,V\,\xi - E\,(\,x-\xi\,)\,\} - m\,]^+ f(\xi)\,\mathrm{d}\xi + \int_x^\infty [-\{\,V\,\xi - U\,(\,\xi - x\,)\,\} - m\,]^+ f(\xi)\,\mathrm{d}\xi. \tag{6.1}$$

Then, we consider three cases so as to evaluate the two integrals in (6.1).

$\langle\langle$ **case 1.** $m < -Vx$ $\rangle\rangle$ Since the net loss $-\mathcal{P}$ is greater than m for any demand ξ in this case (see Figure 6.1), the integral part (6.1) becomes

$$\int_0^x [-\{\,V\,\xi - E\,(\,x-\xi\,)\,\} - m\,]\, f(\xi)\,\mathrm{d}\xi + \int_x^\infty [-\{\,V\,\xi - U\,(\,\xi - x\,)\,\} - m\,]\, f(\xi)\,\mathrm{d}\xi.$$

From the first-order condition of Problem (3.8) , i.e., $\frac{\partial p}{\partial x} = 0$ and $\frac{\partial p}{\partial m} = 0$, we have a solution (x^*, m^*) satisfying $x^* = F^{-1}\big(\frac{U}{E+U}\big)$ and $m^* < -Vx^*$, only when we set $\beta = 0$.

$\langle\langle$ **case 2.** $m \in [-Vx, Ex)$ $\rangle\rangle$ When $\pi > 0$, the integral part (6.1) becomes

$$\int_0^{\frac{Ex-m}{E+V}} [-\{\,V\,\xi - E\,(\,x-\xi\,)\,\} - m\,]\, f(\xi)\,\mathrm{d}\xi$$
$$+ \int_{\frac{Ux+m}{U-V}}^\infty [-\{\,V\,\xi - U\,(\,\xi - x\,)\,\} - m\,]\, f(\xi)\,\mathrm{d}\xi,$$

while the second integral vanishes when $\pi = 0$. From the first-order condition,

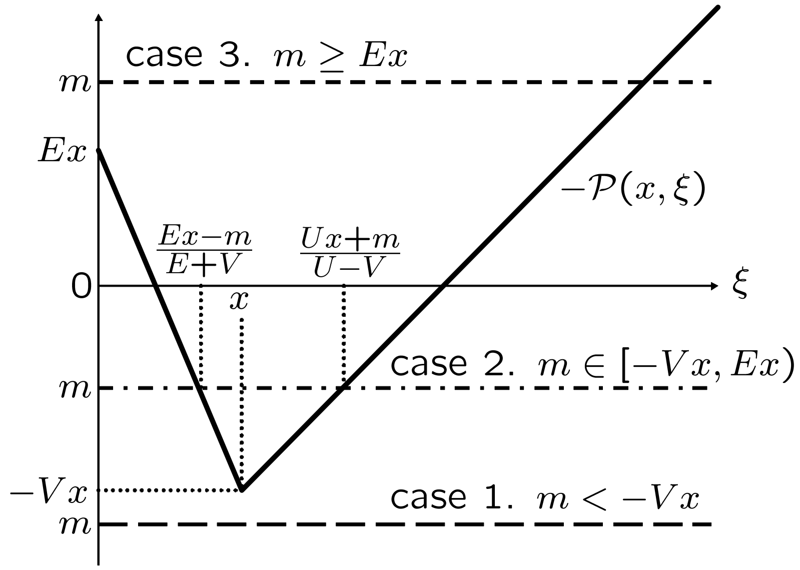


Figure 6.1: Three Cases in Minimization of Net Loss CVaR

we have a solution (x^*, m^*) defined by

$$\begin{cases} x^* = \dfrac{E+V}{E+U}\, F^{-1}\!\left(\dfrac{U(1-\beta)}{E+U}\right) + \dfrac{U-V}{E+U}\, F^{-1}\!\left(\dfrac{E\beta+U}{E+U}\right), \\[2ex] m^* = \dfrac{E\,(U-V)}{E+U}\, F^{-1}\!\left(\dfrac{E\beta+U}{E+U}\right) - \dfrac{U\,(E+V)}{E+U}\, F^{-1}\!\left(\dfrac{U(1-\beta)}{E+U}\right). \end{cases}$$

It is easy to see that this solution (x^*, m^*) satisfies $m^* \in [-Vx^*, Ex^*)$ under Assumption 3.1. Also, we note that this x^* includes the solution in the previous case when $\beta = 0$.

$\langle\langle$ **case 3.** $m \geq Ex\rangle\rangle$ When $\pi = 0$, the integral part (6.1) is shown to be 0, and the problem thus has no solution since any (x, m) does not satisfy the first-order condition. As for the case of $\pi > 0$, the integral (6.1) becomes

$$\int_{\frac{Ux+m}{U-V}}^{\infty} [-\{V\xi - U(\xi - x)\} - m]\, f(\xi)\, \mathrm{d}\xi.$$

By differentiating this equation, we observe that this case also has no optimal solution. ∎

B Proof of Proposition 3.7

As in the proof of Proposition 3.4, we consider the following three cases.

$\langle\langle$ **case 1.** $m < -Ux$ $\rangle\rangle$ Noting that $\pi = 0$ is equivalent to $U = V$, we have the same result as the case 1 of the discussion for deriving the numerical procedure in Section 3.3.1 where $\pi > 0$ is assumed.

$\langle\langle$ **case 2.** $m = -Ux$ $\rangle\rangle$ From the first-order condition, we have a solution (x^*, m^*) defined by

$$x^* = F^{-1}\left(\frac{U}{E+U} \cdot \frac{1+\lambda}{1+\lambda(1-\beta)^{-1}}\right); \quad m^* = -U\,x^*.$$

$\langle\langle$ **case 3.** $m > -Ux$ $\rangle\rangle$ By exploiting Proposition 3.4, we have a solution defined by

$$\begin{cases} x^* = F^{-1}\left(\frac{U-\lambda E}{E+U}\right), \\ m^* = E\,x^* - (\,E+U\,)\,F^{-1}(1-\beta). \end{cases}$$

Combining with the condition $m^* > -Ux^*$, this is valid only for $\lambda < \frac{E+U}{U}\beta - 1$. Since the optimal solution set is convex when a problem is convex, the result follows. ■

C Linearly Approximated Portfolio Dynamics Equations

In order to formulate LAP($\bar{\boldsymbol{u}}$,$\bar{\boldsymbol{w}}$) which is a linear approximation problem for Problem (5.11), the nonlinear terms of decision variables in Constraint (5.9) are linearly approximated at $(\bar{\boldsymbol{u}}, \bar{\boldsymbol{w}})$ with respect to $\boldsymbol{u}$, $\boldsymbol{w}$ as follows:

$$\left\{\begin{aligned}
v_0 &= u_0 + \sum_{i\in\mathcal{N}} \gamma_i(\bar{u}_0\bar{w}_i - p_{0,i}\bar{y}_i) + (u_0 - \bar{u}_0)\sum_{i\in\mathcal{N}} \gamma_i'(\bar{u}_0\bar{w}_i - p_{0,i}\bar{y}_i)\bar{w}_i \\
&\quad + \bar{u}_0 \sum_{i\in\mathcal{N}} \gamma_i'(\bar{u}_0\bar{w}_i - p_{0,i}\bar{y}_i)(w_i - \bar{w}_i), \\
v_1^s &= u_1^s + \sum_{i\in\mathcal{N}} \gamma_i\left(\bar{u}_1^s\bar{w}_i - \frac{p_{1,i}^s\bar{u}_0\bar{w}_i}{p_{0,i}}\right) + (u_1^s - \bar{u}_1^s)\sum_{i\in\mathcal{N}} \gamma_i'\left(\bar{u}_1^s\bar{w}_i - \frac{p_{1,i}^s\bar{u}_0\bar{w}_i}{p_{0,i}}\right)\bar{w}_i \\
&\quad - (u_0 - \bar{u}_0)\sum_{i\in\mathcal{N}} \gamma_i'\left(\bar{u}_1^s\bar{w}_i - \frac{p_{1,i}^s\bar{u}_0\bar{w}_i}{p_{0,i}}\right)\frac{p_{1,i}^s\bar{w}_i}{p_{0,i}} \\
&\quad + \sum_{i\in\mathcal{N}} \gamma_i'\left(\bar{u}_1^s\bar{w}_i - \frac{p_{1,i}^s\bar{u}_0\bar{w}_i}{p_{0,i}}\right)\left(\bar{u}_1^s - \frac{p_{1,i}^s\bar{u}_0}{p_{0,i}}\right)(w_i - \bar{w}_i), \qquad s\in\mathcal{S} \\
v_t^s &= u_t^s + \sum_{i\in\mathcal{N}} \gamma_i\left(\bar{u}_t^s\bar{w}_i - \frac{p_{t,i}^s\bar{u}_{t-1}^s\bar{w}_i}{p_{t-1,i}^s}\right) + (u_t^s - \bar{u}_t^s)\sum_{i\in\mathcal{N}} \gamma_i'\left(\bar{u}_t^s\bar{w}_i - \frac{p_{t,i}^s\bar{u}_{t-1}^s\bar{w}_i}{p_{t-1,i}^s}\right)\bar{w}_i \\
&\quad - (u_{t-1}^s - \bar{u}_{t-1}^s)\sum_{i\in\mathcal{N}} \gamma_i'\left(\bar{u}_t^s\bar{w}_i - \frac{p_{t,i}^s\bar{u}_{t-1}^s\bar{w}_i}{p_{t-1,i}^s}\right)\frac{p_{t,i}^s\bar{w}_i}{p_{t-1,i}^s} \\
&\quad + \sum_{i\in\mathcal{N}} \gamma_i'\left(\bar{u}_t^s\bar{w}_i - \frac{p_{t,i}^s\bar{u}_{t-1}^s\bar{w}_i}{p_{t-1,i}^s}\right)\left(\bar{u}_t^s - \frac{p_{t,i}^s\bar{u}_{t-1}^s}{p_{t-1,i}^s}\right)(w_i - \bar{w}_i), \quad t\in\mathcal{T}\setminus\{1,T\},\ s\in\mathcal{S} \\
v_1^s &= \bar{u}_0\sum_{i\in\mathcal{N}}(1 + r_{1,i}^s)\bar{w}_i + (u_0 - \bar{u}_0)\sum_{i\in\mathcal{N}}(1 + r_{1,i}^s)\bar{w}_i + \bar{u}_0\sum_{i\in\mathcal{N}}(1 + r_{1,i}^s)(w_i - \bar{w}_i), \quad s\in\mathcal{S} \\
v_t^s &= \bar{u}_{t-1}^s\sum_{i\in\mathcal{N}}(1 + r_{t,i}^s)\bar{w}_i + (u_{t-1}^s - \bar{u}_{t-1}^s)\sum_{i\in\mathcal{N}}(1 + r_{t,i}^s)\bar{w}_i \\
&\quad + \bar{u}_{t-1}^s\sum_{i\in\mathcal{N}}(1 + r_{t,i}^s)(w_i - \bar{w}_i), \qquad t\in\mathcal{T}\setminus\{1\},\ s\in\mathcal{S},
\end{aligned}\right. \tag{6.2}$$

where $\gamma_i' : \mathbb{R} \to \mathbb{R}$ $(i \in \mathcal{N})$ represents a subgradient of γ_i, that is

$$\gamma_i'(\eta) := \begin{cases} (d_i\eta + 1)c_i\exp(d_i\eta) & \text{if } \eta > 0 \\ 0 & \text{if } \eta = 0 \\ (b_i\eta - 1)a_i\exp(-b_i\eta) & \text{if } \eta < 0. \end{cases} \tag{6.3}$$

D Convex Subproblem over the Subrectangle

In the rectangular branch-and-bound algorithm for Problem (5.15), the problem over a subrectangle $[\boldsymbol{w}^{\mathcal{L}}, \boldsymbol{w}^{\mathcal{U}}]$ is relaxed by adding the extra quadratic term to the objective function:

$$
\left|
\begin{array}{lll}
\underset{(\tau,\boldsymbol{w})\in\mathbb{R}\times\mathbb{R}^N}{\text{maximize}} & \lambda\left(\dfrac{1}{S}\displaystyle\sum_{s\in\mathcal{S}} v_0 \prod_{t\in\mathcal{T}}\left\{\sum_{i\in\mathcal{N}}(1+r_{t,i}^s)w_i\right\}\right) & \\
& -(1-\lambda)\left(\tau+\dfrac{1}{(1-\beta)S}\displaystyle\sum_{s\in\mathcal{S}}\phi\left(-v_0\prod_{t\in\mathcal{T}}\left\{\sum_{i\in\mathcal{N}}(1+r_{t,i}^s)w_i\right\}-\tau\right)\right) & \\
& -\Theta\displaystyle\sum_{i\in\mathcal{N}}(w_i^{\mathcal{L}}-w_i)(w_i^{\mathcal{U}}-w_i) & \cdots(6.4.\,\text{a}) \\
\text{subject to} & \text{Investment Proportion Constraints (5.10)}, & \\
& w_i^{\mathcal{L}} \le w_i \le w_i^{\mathcal{U}}, & i\in\mathcal{N},
\end{array}
\right. \tag{6.4}
$$

where $\Theta > 0$ is a scalar parameter.

Proposition 6.1 *Function (6.4. a) is concave in $(\tau, \boldsymbol{w})$ for a sufficiently large Θ.* $\star$

Proof : Let $\boldsymbol{H} = (h_{ij}) \in \mathbb{R}^{(1+N)\times(1+N)}$ be a Hessian matrix of "$(-1)\times$(Function (6.4. a))" with respect to $(\tau, \boldsymbol{w})$. To complete the proof it is only necessary to show that $\boldsymbol{H}$ is positive semidefinite. Because $\phi''(\,\cdot\,) > 0$ due to a property of smoothing function [66], we have

$$
\begin{aligned}
h_{11} &= \frac{\partial^2}{\partial\tau\partial\tau}\Big((-1)\times(\text{Function (6.4. a)})\Big) \\
&= \frac{(1-\lambda)}{(1-\beta)S}\sum_{s\in\mathcal{S}}\phi''\left(-v_0\prod_{t\in\mathcal{T}}\left\{\sum_{i\in\mathcal{N}}(1+r_{t,i}^s)w_i\right\}-\tau\right) \ge 0.
\end{aligned}
$$

We show that $(\tau, \boldsymbol{w})^\top \boldsymbol{H}(\tau, \boldsymbol{w}) \ge 0$ for all $(\tau, \boldsymbol{w})$. If $\boldsymbol{w} \ne \boldsymbol{0}$,

$$
(\tau,\boldsymbol{w})^\top\boldsymbol{H}(\tau,\boldsymbol{w}) = 2\Theta\sum_{i=1}^N w_i^2 + \underbrace{\cdots\cdots}_{\Theta\text{ does not appear}},
$$

is nonnegative for a sufficiently large Θ. Otherwise $\boldsymbol{w} = \boldsymbol{0}$, and then we have $(\tau, \boldsymbol{w})^\top \boldsymbol{H}(\tau, \boldsymbol{w}) = \tau^2 h_{11} \ge 0$. ■

Acknowledgments

This paper owes much to the thoughtful and helpful suggestions of Professor Jun-ya Gotoh from Chuo University. I would like to extend my deep appreciation to him. I wish to express my gratitude to my supervisor Professor Yoshitsugu Yamamoto for reading the manuscript and making a number of helpful comments. I am also grateful to Professors Takahito Kuno, Akiko Yoshise, Maiko Shigeno and Masahiro Hachimori for valuable advice. I wish to thank Mizuho-DL Financial Technology Co., Ltd., for generous financial assistance, and Fair Isaac which provides Xpress-MP used in the computational experiments. This work was supported by Grant-in-Aid for JSPS Fellows. Finally, I would like to thank my parents, my friends and members of the team PhilOpt for their kind cooperation.

Bibliography

[1] S. Ahmed, U. Cakmak and A. Shapiro (2007), "Coherent Risk Measures in Inventory Problems," *European Journal of Operational Research*, **182**, 226–238.

[2] M. Anvari (1987), "Optimality Criteria and Risk in Inventory Models: The Case of the Newsboy Problem," *Journal of the Operational Research Society*, **38**, 625–632.

[3] P. Artzner, F. Delbaen, J.-M. Eber and D. Heath (1999), "Coherent Measures of Risk," *Mathematical Finance*, **9**, 203–228.

[4] P. Artzner, F. Delbaen, J.M. Eber and D. Heath (1999), "Coherent Measures of Risk," *Mathematical Finance*, **9**, 203–228.

[5] S. Benati and R. Rizzi (2007), "A Mixed Integer Linear Programming Formulation of the Optimal Mean/Value-at-Risk Portfolio Problem," *European Journal of Operational Research*, **176**, 423–434.

[6] A. Ben-Tal and A. Nemirovski (1999), "Robust Solutions to Uncertain Linear Programs," *Operations Research Letters*, **25**, 1–13.

[7] A. Ben-Tal and A. Nemirovski (2000), "Robust Solutions of Linear Programming Problems Contaminated with Uncertain Data," *Mathematical Programming*, **88**, 411–424.

[8] D. Bertsimas and M. Sim (2004), "Price of Robustness," *Operations Research*, **52**, 35–53.

[9] D. Bertsimas and A. Thiele (2004), "A Data-Driven Approach to Newsvendor Problems," Technical report, Massachusetts Institute of Technology.

[10] D. Bertsimas and A. Thiele (2006), "Robust and Data-Driven Optimization: Modern Decision-Making under Uncertainty," *INFORMS Tutorials in Operations Research*, Chapter 4.

[11] D. Bertsimas and D. Pachamanova (2008), "Robust Multiperiod Portfolio Management in the Presence of Transaction Costs," *Computers & Operations Research*, **35**, 3–17.

[12] J.R. Birge and F. Louveaux (1997), *Introduction to Stochastic Programming*, Springer, New York.

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

[14] M.W. Brandt (1999), "Estimating Portfolio and Consumption Choice: A Conditional Euler Equations Approach," *Journal of Finance*, **54**, 1609–1645.

[15] M.J. Brennan, E.S. Schwartz and R. Lagnado (1997), "Strategic Asset Allocation," *Journal of Economic Dynamics & Control*, **21**, 1377–1403.

[16] G. Calafiore and M.C. Campi (2004), "Decision Making in an Uncertain Environment: The Scenario-Based Optimization Approach," In *Multiple Participant Decision Making*, J. Andrysek, M. Karny and J. Kracik (eds.), Advanced Knowledge International, 99–111.

[17] G. Calafiore and M.C. Campi (2005), "Uncertain Convex Programs: Randomized Solutions and Confidence Levels," *Mathematical Programming*, **102**, 25–46.

[18] G.C. Calafiore (2008), "Multi-Period Portfolio Optimization with Linear Control Policies," *Automatica*, **44**, 2463–2473.

[19] D.R. Cariño, T. Kent, D.H. Myers, C. Stacy, M. Sylvanus, A.L. Turner, K. Watanabe and W.T. Ziemba (1994), "The Russell-Yasuda Kasai Model: An

Asset/Liability Model for a Japanese Insurance Company Using Multistage Stochastic Programming," *Interfaces*, **24**, 29–49.

[20] A. Charnes, W.W. Cooper and G.H. Symonds (1958), "Cost Horizons and Certainty Equivalents: An Approach to Stochastic Programming of Heating Oil," *Management Science*, **4**, 235–263.

[21] S. Choi and A. Ruszczýnski (2008), "A Risk-Averse Newsvendor with Law Invariant Coherent Measures of Risk," *Operations Research Letters*, **36**, 77–82.

[22] G. Cornuejols and R. Tütüncü (2007), *Optimization Methods in Finance*, Cambridge University Press.

[23] T.M. Cover (1991), "Universal Portfolios," *Mathematical Finance*, **1**, 1–29.

[24] G.B. Dantzig (1963), *Linear Programming and Extensions*, Princeton University Press.

[25] G.B. Dantzig and G. Infanger (1993), "Multi-Stage Stochastic Linear Programs for Portfolio Optimization," *Annals of Operations Research*, **45**, 59–76.

[26] V. DeMiguel, L. Garlappi and R. Uppal (2009), "Optimal Versus Naive Diversification: How Inefficient is the $1/N$ Portfolio Strategy?," *Review of Financial Studies*, **22**, 1915–1953.

[27] D. Dentcheva (2006), "Optimization Models with Probabilistic Constraints," In *Probabilistic and Randomized Methods for Design under Uncertainty*, G. Calafiore and F. Dabbene (eds.), Springer, London, 49–97.

[28] L. Eeckhoudt, C. Gollier and H. Schlesinger (1995), " The Risk-Averse (and Prudent) Newsboy," *Management Science*, **41**, 786–794.

[29] L. El Ghaoui, M. Oks, and F. Oustry (2003), "Worst-Case Value-At-Risk and Robust Portfolio Optimization: A Conic Programming Approach," *Operations Research*, **51**, 543–556.

[30] Y. Ermoliev and R.J.-B. Wets (eds.) (1988), *Numerical Techniques for Stochastic Optimization*, Springer-Verlag.

[31] P.C. Fishburn (1977), "Mean-Risk Analysis with Risk Associated with Below Target Returns," *American Economic Review*, **67**, 116–126.

[32] S.-E. Fleten, K. Høyland and S.W. Wallace (2002), "The Performance of Stochastic Dynamic and Fixed Mix Portfolio Models," *European Journal of Operational Research*, **140**, 37–49.

[33] A.A. Gaivoronski and G. Pflug (2005), "Value-at-Risk in Portfolio Optimization: Properties and Computational Approach," *Journal of Risk*, **7**, 1–31.

[34] A.A. Gaivoronski and F. Stella (2003), "On-Line Portfolio Selection Using Stochastic Programming," *Journal of Economic Dynamics & Control*, **27**, 1013–1043.

[35] M. Gilli, E. Këllezi and H. Hysi (2006), "A Data-Driven Optimization Heuristic for Downside Risk Minimization," *Journal of Risk*, **8**, 1–19.

[36] GloriaMundi.org, `http://www.gloriamundi.org/`.

[37] J. Gotoh and Y. Takano (2007), "Newsvendor Solutions via Conditional Value-at-Risk Minimization," *European Journal of Operational Research*, **179**, 80–96.

[38] N. Hibiki (2003), "A Hybrid Simulation/Tree Stochastic Optimisation Model for Dynamic Asset Allocation," In: *Asset and Liability Management Tools: A Handbook for Best Practice*, B. Scherer (eds.), 269–294, Risk Books.

[39] N. Hibiki (2006), "Multi-Period Stochastic Optimization Models for Dynamic Asset Allocation," *Journal of Banking & Finance*, **30**, 365–390.

[40] G. Infanger (2006), "Dynamic Asset Allocation Strategies Using a Stochastic Dynamic Programming Approach," In: *Handbook of Asset and Liability Management*, S.A. Zenios and W.T. Ziemba (eds.), 199–251, Elsevier.

[41] P. Kall and S.W. Wallace (1994), *Stochastic Programming*, John Wiley & Sons Inc.

[42] M. Khouja (1999), "The Single-Period (News-Vendor) Problem: Literature Review and Suggestions for Future Research," *Omega*, **27**, 537–553.

[43] H. Konno and H. Yamazaki (1991), "Mean-Absolute Deviation Portfolio Optimization Model and Its Applications to Tokyo Stock Market," *Management Science* , **37**, 519–531.

[44] H. Konno and S. Suzuki (1992), "A Fast Algorithm for Solving Large Scale Mean-Vaviance Models by Compact Factorization of Covariance Matrices," *Journal of the Operations Research Society of Japan*, **35**, 93–104.

[45] H. Konno, P.T. Thach and H. Tuy (1997), *Optimization on Low Rank Nonconvex Structures*, Kluwer Academic, Dordrecht.

[46] J. Kottas and H. Lau (1980), "The Use of Versatile Distribution Families in Some Stochastic Inventory Calculations," *Journal of the Operational Research Society*, **31**, 393–403.

[47] M.I. Kusy and W.T. Ziemba (1986), "A Bank Asset and Liability Management Model," *Operations Research*, **34**, 356–376.

[48] N. Larsen, H. Mausser and S. Uryasev (2002), "Algorithms for Optimization of Value-at-Risk," *Financial Engineering, e-Commerce and Supply Chain*, P. Pardalos and V. K. Tsitsiringos (Eds.), Kluwer Academic Publishers, 129–157.

[49] A. Lau and H. Lau (1988), "Maximizing the Probability of Achieving a Target Profit Level in a Two-Product Newsboy Problem," *Decision Sciences*, **19**, 392–408.

[50] H. Lau (1980), "The Newsboy Problem under Alternative Optimization Objectives," *Journal of the Operational Research Society*, **31**, 525–535.

[51] C.D. Maranas and C.A. Floudas (1994), "Global Minimum Potential Energy Conformations of Small Molecules," *Journal of Global Optimization*, **4**, 135–170.

[52] C.D. Maranas, I.P. Androulakis, C.A. Floudas, A.J. Berger and J.M. Mulvey (1997), "Solving Long-Term Financial Planning Problems via Global Optimization," *Journal of Economic Dynamics & Control*, **21**, 1405–1425.

[53] H.M. Markowitz (1952), "Portfolio Selection," *Journal of Finance*, **7**, 77–91.

[54] H. Markowitz (1991), *Portfolio Selection: Efficient Diversification of Investments (2nd edition)*, Wiley, New York.

[55] R.C. Merton (1969), "Lifetime Portfolio Selection under Uncertainty: The Continuous-Time Case," *Review of Economics and Statistics*, **51**, 247–257.

[56] R.C. Merton (1971), "Optimum Consumption and Portfolio Rules in a Continuous-Time Model," *Journal of Economic Theory*, **3**, 373–413.

[57] L.B. Miller and H. Wagner (1965), "Chance Constrained Programming with Joint Constraints," *Operations Research*, **13**, 930–945.

[58] J.M. Mulvey and H. Vladimirou (1989), "Stochastic Network Optimization Models for Investment Planning," *Annals of Operations Research*, **20**, 187–217.

[59] J.M. Mulvey and W.T. Ziemba (1995), "Asset and Liability Allocation in a Global Environment," In: *Finance (Handbooks in Operations Research and Management Science, Vol. 9)*, R.A. Jarrow, V. Maksimovic and W.T. Ziemba (eds.), 435–463, North-Holland.

[60] K. Natarajan, M. Sim and J. Uichanco (2009), "Asymmetry and Ambiguity in Newsvendor Models," Working Paper, National University of Singapore.

[61] K. Natarajan, D. Pachamanova and M. Sim (2008), "Incorporating Asymmetric Distributional Information in Robust Value-at-Risk Optimization," *Management Science*, **54**, 573–585.

[62] A. Nemirovski (2003), "On Tractable Approximations of Randomly Perturbed Convex Constraints," *Proceedings of the 42nd IEEE Conference on Decision and Control*, **3**, 2419–2422.

[63] A. Nemirovski and A. Shapiro (2006), "Convex Approximations of Chance Constrained Programs," *SIAM Journal on Optimization*, **17**, 969–996.

[64] W. Ogryczak and A. Ruszczýnski (2001), "On Consistency of Stochastic Dominance and Mean-Semideviation Models," *Mathematical Programming* Ser.B, **89**, 217–232.

[65] W. Ogryczak and A. Ruszczýnski (2002), “Dual Stochastic Dominance and Related Mean-Risk Models,” *SIAM Journal on Optimization*, **13**, 60–78.

[66] J.-S. Pang and S. Leyffer (2004), “On the Global Minimization of the Value-at-Risk,” *Optimization Methods and Software*, **19**, 611–631.

[67] M. Parlar and Z.K. Weng (2003), “Balancing Desirable but Conflicting Objectives in the Newsvendor Problem,” *IIE Transactions*, **35**, 131–142.

[68] A.F. Perold and W.F. Sharpe (1995), “Dynamic Strategies for Asset Allocation,” *Financial Analysts Journal*, **51**, 149–160.

[69] G.Ch. Pflug (2000), “Some Remarks on the Value-at-Risk and the Conditional Value-at-Risk,” In: *Probabilistic Constrained Optimization: Methodology and Applications*, S. Uryasev (eds.), 272–281, Kluwer Academic Publishers, Dordrecht.

[70] M.Ç. Pinar (2007), “Robust Scenario Optimization Based on Downside-Risk Measure for Multi-Period Portfolio Selection,” *OR Spectrum*, **29**, 295–309.

[71] A. Prékopa (1970), On Probabilistic Constrained Programming, *Proceedings of the Princeton Symposium on Mathematical Programming*, Princeton University Press, Princeton, 113–138.

[72] A.v. Puelz (2001), “Value-at-Risk Based Portfolio Optimization,” *Stochastic Optimization: Algorithms and Applications*, S. Uryasev and P. M. Pardalos (Eds.), Kluwer Academic Publishers, 279–302.

[73] R.T. Rockafellar (1970), *Convex Analysis*, Princeton University Press, NJ.

[74] R.T. Rockafellar and S. Uryasev (2002), “Conditional Value-at-Risk for General Loss Distributions,” *Journal of Banking & Finance*, **26**, 1443–1471.

[75] P.A. Samuelson (1969), “Lifetime Portfolio Selection by Dynamic Stochastic Programming,” *Review of Economics and Statistics*, **51**, 239–246.

[76] A. Shapiro, D. Dentcheva and A. Ruszczýnski (2009), *Lectures on Stochastic Programming: Modeling and Theory*, SIAM, PA.

[77] E.A. Silver, D.F. Pyke and R. Peterson (1998), *Inventory Management and Production Planning and Scheduling 3rd Edition*, Wiley, NJ.

[78] J. Skaf and S. Boyd (2008), "Multi-Period Portfolio Optimization with Constraints and Transaction Costs," Working Paper, Stanford University.

[79] Y. Takano and J. Gotoh, "α-Conservative Approximation for Probabilistically Constrained Convex Programs," *Computational Optimization and Applications*, to appear.

[80] Y. Takano and J. Gotoh (2009), "Constant Rebalanced Portfolio Optimization under Nonlinear Transaction Costs," Discussion Paper, University of Tsukuba.

[81] H. Tuy (1998), *Convex Analysis and Global Optimization*, Kluwer Academic Publishers, Dordrecht.

[82] A. Verma (2005), "VaR optimal portfolios. A Global Optimization Approach," Workshop on Optimization in Finance, Coimbra.

[83] M.R. Young (1998), "A Minimax Portfolio Selection Rule with Linear Programming Solution," *Management Science*, **44**, 673–683.

[84] X.-L. Zhang, K.-C. Zhang (2009), "Using Genetic Algorithm to Solve a New Multi-Period Stochastic Optimization Model," *Journal of Computational and Applied Mathematics*, **231**, 114–123.

[85] W.T. Ziemba and J.M. Mulvey (eds.) (1998), *Worldwide Asset and Liability Modeling*, Cambridge University Press.