Abstract

In general, prices of financial assets, such as stocks, fluctuate uncertainly. Most of the existing portfolio optimization models assume that the random processes of the prices are known in advance. Thus, we need to estimate the processes before formulating the models. We usually use historical data of the assets for the estimation, and hence the portfolio obtained from the model with the estimation deeply depends on the specified situation where the data are picked out. Thus the portfolio may not be appropriate to other situations. Moreover, when we do not have sufficient data that match the current situation, we cannot have a good portfolio for the situation. If we can change the portfolio according to the economical situations at the decision times of the investment, we can expect to control the risk and the return of the assets more flexibly. Thus, we attempt to construct a function that outputs the optimal portfolio of the current economical situation by directly using the long-time historical data of the assets.

In this paper, we call such a function a portfolio function. By replacing a decision variable in the traditional portfolio optimization models with a portfolio function, we easily formulate optimization models whose solution is the optimal portfolio function. However, since the decision variable of such optimization model is a function, the models are formulated as infinite programming problems which are not easily solved. To overcome the difficulty, we restrict the portfolio function to linear combinations of certain basis functions. Moreover, we directly exploit the historical data to approximate the random behaviors of the assets. Thereby, the infinite problem is reduced to the convex quadratic programming problem. We also consider its dual problem, and show that the dual problem is formulated by using some kernel functions. This so-called kernel method provides the optimal portfolio function represented with the kernel functions only. Unfortunately the dual problem become very large and dense, and hence we cannot apply the standard solvers, such as the interior point method. Thus, we propose to apply the matrix splitting method that exploits the special structure of the dual problem. We present some results of numerical experiments with practical data, and discuss the validity of the proposed approach.

Contents

1 Introduction

In general, prices of financial assets, such as stocks, fluctuate uncertainly. Hence, when investors allocate their funds in the assets, they have to consider not only the return of the investment but also the risk. Most of them invest in various assets to keep a certain amount of the return, and to make the risk as small as possible. This allocation is called a portfolio, and optimization problems which yield an optimal portfolio is called a portfolio optimization model.

Until now various portfolio optimization models have been proposed [4, 5, 7, 9, 10]. One of the most important task for the modeling is to decide how to represent a risk of portfolios. In the Markowitz's portfolio optimization model [7], called mean-variance model (MV model), the risk is defined as the variance of the return rate of the portfolio. The model is to minimize the variance while keeping the expected return rate above a certain level. It can be formulated as the convex quadratic programming problem (convex QP), and hence we can solve the problem by the interior point method [8]. Konno and Yamazaki [5] proposed the mean-absolute deviation model (MAD model) which adopts the expected absolute deviation of the return rate as the risk of the portfolio. By using the historical data of the assets, it can be formulated as the linear programming problem. Hence, the MAD model is much easier to solve than the MV model. The above two models that adopt the variance or the absolute deviation of the return rate as the risk are supposed that the investors consider a return larger than they expected as undesirable. However, the larger return is usually favorable. Therefore Konno, Waki and Yuuki [4] considered the lower semi-variance (LSV) and the lower semi-absolute deviation (LSAD) as the more practical risks. The well-known Value-at-Risk (VaR) is also one such of risk measures. Given the constant $\beta \in (0, 1)$, the $\beta-$VaR is defined as the minimum loss of the worst loss happening at the probability of $1 - \beta$. By the definition, the loss is less than the $\beta-$VaR at the probability of β. The VaR is very easily understandable as a risk, and hence it is popular to evaluate various financial assets. However, the minimization problem of the VaR minimization cannot be formulated as the convex programming problem. Thus it is hard to obtain the optimal portfolio based on VaR minimization [10]. Moreover, it is known that the VaR does not satisfy the important properties of the risk measure, such as coherent. To overcome the difficulty, the conditional Value-at-Risk (CVaR) have been proposed [10]. The CVaR is defined as the expectation of the loss over the VaR. The portfolio optimization models with the CVaR is formulated as the convex programming problem.

In order to formulate the above popular models, we have to estimate the random behaviors of the prices of the assets in advance. The estimation is one of the most important topics in the financial engineering research. One of the popular methods for the estimation is to use the realizations of the random variables obtained directly from the historical data [7]. This method is very simple and easy to implement. However, its validity deeply depends on the choice of the appropriate historical data. Sharpe [12], Markowitz and Perold [9] proposed the factor model that use some economical factors to estimate a price of each asset. The model is assumed that the return of each asset is represented by some factors such as, exchange rates and industry indices. Markowitz and Perold [9] formulated the MV model with the factor model. The scenario based estimation [6] is also useful. In this method we prepare multiple scenarios which should occur in the future, and set a occurrence probability of each scenario. Then we compute the risk and the expected return in the scenarios. It is also important for the scenario based

method to estimate the occurrence probability of each scenario.

As described above, we should first estimate the probability distribution of the price of the assets. Then we formulate a portfolio optimization model based on the estimation. Finally we get the desired portfolio by solving the formulated model. Therefore, the obtained portfolio would be optimal for the situation where the probability distribution is estimated. However, it might not be appropriate to the different situations. To obtain an optimal portfolio for the specific situation, we have to estimate the probability distribution under the situation. In general, to estimate the probability distribution accurately, we must have a sufficiently-large number of data. Hence, if the number of data for the situation is not sufficient, the estimated probability distribution become less-accurate. However, the number of data is finite, hence we may not be able to have sufficient data corresponding to each individual situation. For example, there may not exist the data in the case where the exchange rate is 101 yen to the dollar. Even in such a case, there may exist sufficient data for 100 and 102 yens. Then we would estimate the distribution under 101 yen by using the data, and get the optimal portfolio for 101 yen. However, this process requires for us to choose the appropriate data (100 and 102 yen), which is not a easy task. It would be desirable to choose data automatically or to give a function that outputs the optimal portfolio according to the current situation. Therefore, in this paper, we attempt to construct such a function.

We call the function a portfolio function. We can formulate the optimization models for the portfolio function by naturally extending the existing portfolio optimization models. However, since the decision variables of the extended models are functions, the extended models become the infinite programming problem which are not easily solved. Thus, we propose to restrict the portfolio function to the linear combinations of some feature vector functions, and to use the historical data. Thereby, we can reduce the infinite programming problem to the convex QP. The number of the decision variables of the convex QP is equivalent to the dimension of the feature vector function. In general, high dimension is desirable to describe the portfolio function flexibly. However, the larger the dimension is, the larger the scale of the convex QP is. Thus, we consider the Lagrangian dual problem of the convex QP. The number of the decision variable of the dual problem is in proportion to the number of assets and the periods of historical data, and is not related to the dimension of the feature vector function. Moreover, we can represent the portfolio function as the linear combinations of kernel functions without computing the feature vector explicitly.

However, even if the model is represented as the finite model, the dual problem becomes very large. For example, when the number of the objective assets and the periods of the historical data is 500 and 1000, respectively, the number of the decision variables is over 500000 (see Section 4 in the detail). In such a large-scale problem, it is difficult to apply the standard solver of the convex QPs such as the interior point method. One of the methods solving very large-scale problems is the matrix splitting method [3]. The matrix splitting method is originally used to solve the systems of linear equations, and it is applied to solve the linear complementarity problems, the mixed complementarity problems (MCPs) and the general convex QPs. The matrix splitting method needs some restricted assumption for global convergence. Moreover, if it cannot be implemental for parallel computers, it takes long time for convergence. Thus it is not popular in general. Fortunately, the Hessian of the objective function of the dual problem has a special structure which satisfies the assumption for global convergence and which is separable for parallel computing. Hence we consider the matrix splitting method for the dual problem.

This paper is organized as follows: In Section 2, we recall the existing models, that is, the mean-LSAD

model and the CVaR model. Moreover, we describe the matrix splitting method for the MCP. In Section 3, we propose the portfolio function optimization models with the LSAD and CVaR minimizations. In Section 4, we derive the Lagrangian dual problems of these models, and we apply the kernel methods to represent an optimal portfolio function by the linear combinations of some kernel functions. In Section 5, we reformulate the dual problem as the MCP, and then we describe how to implement the matrix splitting method for the MCP. In Section 6, we report some numerical results of the proposed approach. Finally, Section 7 concludes the paper.

2 Preliminaries

In this section, we introduce the mean-lower semi-absolute deviation (M-LSAD) model [4] and the CVaR minimization model [10]. We also describe the matrix splitting method for MCPs [3].

2.1 The Existing Portfolio Optimization Models

We first give the formulations of the portfolio optimization models used in subsequent sections. Let $S_i,\ (i = 1, 2, \ldots, n)$ be the objective assets. Also let y_i be the rate of the fund to be allocated to S_i, and let $\boldsymbol{y} = (y_1, y_2, \ldots, y_n)$ be the portfolio which an investor has. In this paper, we do not consider the short selling. Then, the portfolio $\boldsymbol{y}$ must satisfy the following conditions:

$$\sum_{i=1}^{n} y_i \leq 1, y_i \geq 0, i = 1, 2, \ldots, n.$$

Note that there exist uninvested funds in cash at the rate of $1 - \sum_{i=1}^{n} y_i$. Let R_i be the return rate (per period) of the asset S_i. Note that R_i are random variables. The return rate $R(\boldsymbol{y})$ of the portfolio $\boldsymbol{y}$ is represented as

$$R(y) = \sum_{i=1}^{n} R_i y_i.$$

The return $R(\boldsymbol{y})$ is also a random variable. Let $\mathrm{V}(\boldsymbol{y})$ be a risk measure of the portfolio $\boldsymbol{y}$, and let γ be the minimal return rate required by the investor.

Now we discuss the following model:

$$\begin{aligned} \text{minimize} \quad & \mathrm{V}(\boldsymbol{y}) \\ \text{subject to} \quad & y_i \geq 0, \quad i = 1, 2, \ldots, n, \\ & \sum_{i=1}^{n} y_i \leq 1, \\ & \sum_{i=1}^{n} E[R_i] y_i \geq \gamma, \end{aligned} \tag{2.1}$$

where $E[\cdot]$ denotes the expectation of the random variable.

In the portfolio optimization models, the definition of the risk measure $\mathrm{V}(\boldsymbol{y})$ is important. Until now various risk measures are proposed. Now we introduce LSAD and CVaR as the risk.

Konno and Yamazaki [5] proposed the mean-absolute deviation model (MAD model). In the MAD model, the risk is represented as the following measure V_a, called the absolute deviation:

$$\mathrm{V}_a(\boldsymbol{y}) = E\left[\left| \sum_{i=1}^{n} R_i y_i - E\left[\sum_{i=1}^{n} R_i y_i\right] \right|\right].$$

The advantage of the MAD model is that it is formulated as the linear programming problem, while the MV model [7] is formulated as the QP.

In the MAD model, the rate larger than the expected return rate $\sum_{i=1}^{n} E[R_i]y_i$ is also considered as the risk. Thus, this model is less expressing to the investor's feeling. The lower semi-absolute deviation (LSAD) of the return rate $R(\boldsymbol{y})$, denoted by V_l, is a risk measure that regard only the short of the expected return goal as unfavorable.

$$\mathrm{V}_l(\boldsymbol{y}) = E[\max\{0, \alpha - R(\boldsymbol{y})\}],$$

where α is the goal of the return rate per period. For example, we can choose the expected return of $R(\boldsymbol{y})$ or the return of the benchmark as α. Note that the former is a constant, and the latter is a random variable. The portfolio optimization problem (2.1) with the LSAD is called the mean-LSAD (M-LSAD) model.

In general, since the exact probability distribution of R_i is unknown, it is difficult to compute the expectation of $R(\boldsymbol{y})$ and the LSAD. Therefore we often use the estimated values by the historical data as follows. Let s_i^t be the price of the asset S_i during period t, and let r_i^t be the realization of random variable R_i during the period t $(t = 1, 2, \ldots, T)$. Then r_i^t is given by

$$r_i^t = \frac{s_i^{t+1} - s_i^t}{s_i^t}. \tag{2.2}$$

By using the realization r_i^t, the expectation of random variable R_i and the LSAD can be approximated as

$$E[R_i] \approx \frac{1}{T}\sum_{t=1}^{T} r_i^t,$$

$$\mathrm{V}_l(\boldsymbol{y}) \approx \frac{1}{T}\sum_{t=1}^{T} \max\left\{0, \alpha^t - \sum_{i=1}^{n} r_i^t y_i\right\},$$

where α^t is the realization of α during period t. If α is a constant, then $\alpha^t = \alpha$ for all t.

By using the above approximations, the M-LSAD model is represented as follows:

$$\begin{aligned}
\text{minimize} \quad & \frac{1}{T}\sum_{t=1}^{T} \max\left\{0, \alpha^t - \sum_{i=1}^{n} r_i^t y_i\right\} \\
\text{subject to} \quad & y_i \geq 0, \quad i = 1, 2, \ldots, n, \\
& \sum_{i=1}^{n} y_i \leq 1, \\
& \frac{1}{T}\sum_{t=1}^{T}\sum_{i=1}^{n} r_i^t y_i \geq \gamma.
\end{aligned}$$

Introducing an auxiliary variable $\boldsymbol{\eta} \in R^T$, this problem becomes equivalent to the linear programming

problem

$$\begin{aligned}
\text{minimize} \quad & \frac{1}{T}\sum_{t=1}^{T}\eta^t \\
\text{subject to} \quad & \eta^t \geq \alpha^t - \sum_{i=1}^{n} r_i^t y_i, \quad t = 1, 2, \ldots, T, \\
& \eta^t \geq 0, \quad t = 1, 2, \ldots, T, \\
& y_i \geq 0, \quad i = 1, 2, \ldots, n, \\
& \sum_{i=1}^{n} y_i \leq 1, \\
& \frac{1}{T}\sum_{t=1}^{T}\sum_{i=1}^{n} r_i^t y_i \geq \gamma.
\end{aligned} \tag{2.3}$$

The portfolio $\boldsymbol{y}$ obtained by solving the problem (2.3) depends on the data set $\{r_i^t\}$. Therefore, to obtain the portfolio $\boldsymbol{y}$ which is suitable for the particular economical situation, we have to collect the data set $\{r_i^t\}$ fitting the situation.

Next we introduce the CVaR model. Let $f(\boldsymbol{y}, \boldsymbol{R})$ be the loss function for the portfolio $\boldsymbol{y}$. One of the popular loss function is $f(\boldsymbol{y}, \boldsymbol{R}) = -\sum_{i=1}^{n} R_i y_i$. In what follows, we only consider $f(\boldsymbol{y}, \boldsymbol{R}) = -\sum_{i=1}^{n} R_i y_i$ as the loss function.

Since the loss $f(\boldsymbol{y}, \boldsymbol{R})$ depends on the random variable $\boldsymbol{R}$, the loss is also a random variable. Let $\Phi(\cdot|\boldsymbol{y})$ be the distribution function of the loss, that is,

$$\Phi(v|\boldsymbol{y}) = \int_{f(\boldsymbol{y},\boldsymbol{R})\leq v} p(\boldsymbol{R}) d\boldsymbol{R},$$

where $p(\cdot)$ denotes the probability density function of $\boldsymbol{R}$.

Note that $\Phi(v|\boldsymbol{y}) \geq \beta$ means that the probability that the loss becomes over v is $1-\beta$ or less. The β-VaR, denoted by $\mathrm{V}_{\mathrm{VaR}}(\boldsymbol{y}; \beta)$, is defined as the minimum loss such that the value of distribution function $\Phi(v|\boldsymbol{y})$ is over β.

$$\mathrm{V}_{\mathrm{VaR}}(\boldsymbol{y}; \beta) = \min\{v \mid \Phi(v|\boldsymbol{y}) \geq \beta\}.$$

The β-VaR $\mathrm{V}_{\mathrm{VaR}}(\boldsymbol{y}; \beta)$ is not convex with respect to $\boldsymbol{y}$. Then Rockafellar and Uryasev [10] proposed the CVaR as the risk, related to VaR. The CVaR, denoted by $\mathrm{V}_{\mathrm{CVaR}}(\boldsymbol{y}; \beta)$, is defined as the expectation of the loss over the β-VaR.

$$\mathrm{V}_{\mathrm{CVaR}}(\boldsymbol{y}; \beta) = \mathrm{V}_{\mathrm{VaR}}(\boldsymbol{y}; \beta) + \frac{1}{1-\beta} E\left[\max\{0, f(\boldsymbol{y}, \boldsymbol{R}) - \mathrm{V}_{\mathrm{VaR}}(\boldsymbol{y}; \beta)\}\right].$$

Moreover, Rockafellar and Uryasev [10] shows that the CVaR is given by

$$\mathrm{V}_{\mathrm{CVaR}}(\boldsymbol{y}; \beta) = \min_{v} F(\boldsymbol{y}, v; \beta),$$

where

$$F(\boldsymbol{y}, v; \beta) = v + \frac{1}{1-\beta} E\left[\max\{0, f(\boldsymbol{y}, \boldsymbol{R}) - v\}\right].$$

Therefore we can obtain the CVaR without computing the VaR. Moreover, the CVaR minimization

problem is represented as follows:

$$\begin{aligned} \text{minimize}_{v,\boldsymbol{y}} \ & v + \frac{1}{1-\beta} E\left[\max\{0, f(\boldsymbol{y}, \boldsymbol{R}) - v\}\right] \\ \text{subject to} \quad & y_i \geq 0, \quad i = 1, 2, \ldots, n, \\ & \sum_{i=1}^{n} y_i \leq 1, \\ & \frac{1}{T} \sum_{t=1}^{T} \sum_{i=1}^{n} E[R_i y_i] \geq \gamma, \end{aligned} \tag{2.4}$$

where the decision variables are $\boldsymbol{y}$ and v.

When the probability distribution of R_i is unknown, it is difficult to compute the CVaR. Then, in a way similar to the MAD model, we estimate the CVaR by using the historical data [10]. Using r_i^t defined by (2.2), $F(\boldsymbol{y}, v; \beta)$ is approximated as

$$F(\boldsymbol{y}, v; \beta) = v + \frac{1}{(1-\beta)T} \sum_{t=1}^{T} \max\{f(\boldsymbol{y}, \boldsymbol{r^t}) - v\},$$

where $\boldsymbol{r^t} = (r_1^t, r_2^t, \ldots, r_n^t)^\top$. By this approximation, the CVaR minimization model is reduce to

$$\begin{aligned} \text{minimize}_{v,\boldsymbol{y}} \quad & v + \frac{1}{(1-\beta)T} \sum_{t=1}^{T} \max\{0, f(\boldsymbol{y}, \boldsymbol{r^t}) - v\} \\ \text{subject to} \quad & y_i \geq 0, \quad i = 1, 2, \ldots, n, \\ & \sum_{i=1}^{n} y_i \leq 1, \\ & \frac{1}{T} \sum_{t=1}^{T} \sum_{i=1}^{n} r_i^t y_i \geq \gamma, \end{aligned}$$

which is equivalent to the following linear programming problem:

$$\begin{aligned} \text{minimize}_{v,\boldsymbol{y},\eta} \ & v + \frac{1}{(1-\beta)T} \sum_{t=1}^{T} \eta^t \\ \text{subject to} \quad & \eta^t \geq -\sum_{i=1}^{n} r_i^t y_i - v, \quad t = 1, 2, \ldots, T, \\ & \eta^t \geq 0, \quad t = 1, 2, \ldots, T, \\ & y_i \geq 0, \quad i = 1, 2, \ldots, n, \\ & \sum_{i=1}^{n} y_i \leq 1, \\ & \sum_{i=1}^{n} r_i^t y_i \geq \gamma. \end{aligned} \tag{2.5}$$

2.2 The Matrix Splitting Methods for MCP

The portfolio optimization models proposed in Section 4 become the convex QPs. The convex QP is reformulated as the mixed complementarity problem (MCP). In this subsection, we describe the matrix splitting method to solve large-scale MCPs.

First we define a MCP. Given $\boldsymbol{l},\ \boldsymbol{u} \in R^n$ with $l_i < u_i$ $(i = 1, 2, \ldots, n)$ and a mapping $F : R^n \to R^n$, MCP$(F, \boldsymbol{l}, \boldsymbol{u})$ is defined by

$$\begin{array}{lll} \text{Find} & \boldsymbol{z} \in S & \\ \text{such that} & z_i = l_i & \Rightarrow F_i(\boldsymbol{z}) \geq 0, \\ & l_i < z_i < u_i & \Rightarrow F_i(\boldsymbol{z}) = 0, \\ & z_i = u_i & \Rightarrow F_i(\boldsymbol{z}) \leq 0, \end{array} \tag{2.6}$$

where $S = \{\boldsymbol{z} \mid l_i \leq z_i \leq u_i\}$. We allow l_i and u_i to be $l_i = -\infty$ and $u_i = +\infty$, respectively. In this paper, we assume that F is linear. When $F(\boldsymbol{z}) = M\boldsymbol{z} + \boldsymbol{d}$, we write MCP$(F, \boldsymbol{l}, \boldsymbol{u})$ as MCP$(M, d, \boldsymbol{l}, \boldsymbol{u})$. When $l_i = -\infty$ and $u_i = +\infty$ for all i, MCP$(F, \boldsymbol{l}, \boldsymbol{u})$ is reduced to a system of linear equations $F(\boldsymbol{z}) = 0$. Moreover, when $l_i = 0$ and $u_i = +\infty$ for all i, MCP$(F, \boldsymbol{l}, \boldsymbol{u})$ becomes the following linear complementarity problem:

$$\begin{array}{ll} \text{Find} & \boldsymbol{z} \in R^n_+ \\ \text{such that} & F_i(\boldsymbol{z}) \geq 0, \\ & z_i > 0 \Rightarrow F_i(\boldsymbol{z}) = 0, \end{array}$$

where $R^n_+ = \{\boldsymbol{z} \mid z_i \geq 0\ (i = 1, 2, \ldots, n)\}$.

The general convex QP can be reformulated as MCP$(F, \boldsymbol{l}, \boldsymbol{u})$. In what follows, we consider the following convex QP:

$$\begin{array}{ll} \text{minimize} & \dfrac{1}{2}\boldsymbol{z}^\top Q\boldsymbol{z} + \boldsymbol{d}^\top \boldsymbol{z} \\ \text{subject to} & \boldsymbol{l} \leq \boldsymbol{z} \leq \boldsymbol{u}, \\ & A\boldsymbol{z} = \boldsymbol{b}, \end{array} \tag{2.7}$$

where $Q \in R^{n\times n}, A \in R^{m\times n}, \boldsymbol{d} \in R^n$ and $\boldsymbol{b} \in R^m$. The Karush-Kuhn-Tucker conditions (KKT conditions in short) of problem (2.7) is written as

$$\begin{array}{l} Qz + \boldsymbol{d} - \boldsymbol{\lambda_1} + \boldsymbol{\lambda_2} + A^\top \boldsymbol{\lambda_3} = 0, \\ \boldsymbol{z} \geq \boldsymbol{l}, \boldsymbol{\lambda_1} \geq 0, \boldsymbol{\lambda_1}^\top \boldsymbol{z} = 0, \\ \boldsymbol{z} \leq \boldsymbol{u}, \boldsymbol{\lambda_2} \geq 0, \boldsymbol{\lambda_2}^\top (\boldsymbol{u} - \boldsymbol{z}) = 0, \\ A\boldsymbol{z} - \boldsymbol{b} = 0, \end{array} \tag{2.8}$$

where $\boldsymbol{\lambda_1}, \boldsymbol{\lambda_2}, \boldsymbol{\lambda_3}$ are the Lagrange multipliers corresponding to the constraints. The KKT conditions (2.8) are written as MCP$(F, \boldsymbol{l}, \boldsymbol{u})$, where $F, \boldsymbol{l}, \boldsymbol{u}$ are given by

$$\begin{array}{l} F(\boldsymbol{z}, \boldsymbol{\lambda_3}) = \begin{pmatrix} Q\boldsymbol{z} + \boldsymbol{d} + A^\top \boldsymbol{\lambda_3} \\ A\boldsymbol{z} - \boldsymbol{b} \end{pmatrix}, \\ \boldsymbol{l} = (\underbrace{0, \ldots, 0}_{n}, \underbrace{-\infty, \ldots, -\infty}_{m})^\top, \\ \boldsymbol{u} = (\underbrace{u_1, \ldots, u_n}_{n}, \underbrace{+\infty, \ldots, +\infty}_{m})^\top, \end{array} \tag{2.9}$$

respectively.

We next explain the matrix splitting method for MCP$(F, \boldsymbol{l}, \boldsymbol{u})$. Usually, one of the standard solvers of the convex QPs is the interior point method. However, when the problem is large scale and the matrix Q is dense, the interior point method takes much time to solve the problem. When Q has a special structure

and the matrix splitting method can exploit the structure, the matrix splitting method is useful for the problem. As mentioned in Section 5, our problem has such a structure.

We now suppose that F is represented as $F(\boldsymbol{z}) = M\boldsymbol{z} + \boldsymbol{d}$.

We call a pare of matrices such that $M = B + C$ a splitting of the matrix M. The matrix splitting method for MCP$(M, \boldsymbol{d}, \boldsymbol{l}, \boldsymbol{u})$ generates a sequence $\{\boldsymbol{z}^{(k)}\}$ by setting a solution of MCP$(B, C\boldsymbol{z}^{(k)} + \boldsymbol{d}, \boldsymbol{l}, \boldsymbol{u})$ as $\boldsymbol{z}^{(k+1)}$.

Now we explain popular splittings. Let D, L and U be given by

$$D = \begin{pmatrix} M_{11} & & & \\ & M_{22} & & \\ & & \ddots & \\ & & & M_{NN} \end{pmatrix}, \quad L = \begin{pmatrix} 0 & & & \\ M_{21} & \ddots & & \\ \vdots & \ddots & \ddots & \\ M_{N1} & \cdots & M_{N(N-1)} & 0 \end{pmatrix}, \quad U = \begin{pmatrix} 0 & M_{12} & \cdots & M_{1N} \\ & \ddots & \ddots & \vdots \\ & & \ddots & M_{(N-1)N} \\ & & & 0 \end{pmatrix},$$

where M_{ij} $(i, j = 1, 2, \ldots, N)$ are $n_i \times n_j$ submatrices of M. Note that n_i is the constant such that $\sum_{i=1}^{N} n_i = n$. Using these matrices, we can consider the following two types of splittings (B, C) [3].

$$\text{(The block Jacobi)} \quad M = D, \qquad N = L + U. \tag{2.10}$$

$$\text{(The block SOR)} \quad M = \frac{1}{\zeta}D + L, \quad N = \left(1 - \frac{1}{\zeta}\right)D + U. \tag{2.11}$$

The SOR stands for successive overrelaxation. In particular, we call the block-SOR method with $\zeta = 1$ the block-Gauss-Seidel method. When we use the block-SOR method (2.11), we can split the original subproblem MCP$(B, C\boldsymbol{z}^{(k)} + \boldsymbol{d}, \boldsymbol{l}, \boldsymbol{u})$ as $\boldsymbol{z}^{(k+1)}$ in the following N smaller MCPs.

$$\text{MCP}(\zeta^{-1}M_{ii},\ \sum_{j<i} M_{ij}\boldsymbol{z}_j^{(k+1)} + \sum_{j\geq i} M_{ij}\boldsymbol{z}_j^{(k)} - \zeta^{-1}M_{ii}\boldsymbol{z}_i^{(k)} + \boldsymbol{d}_i,\ \boldsymbol{l}_i,\ \boldsymbol{u}_i),\ i = 1, 2, \ldots, N, \tag{2.12}$$

where $\boldsymbol{z}_i, \boldsymbol{l}_i$ and $\boldsymbol{u}_i$ are the corresponding subvectors of $\boldsymbol{z}, \boldsymbol{l}$ and $\boldsymbol{u}$, respectively. Moreover, if we set $N_i = 1, N_j = 1 (i, j = 1, \ldots, N)$, the procedure (2.12) can be simply written as

$$z_i^{(k+1)} = \text{mid}\{l_i, z_i^{(k)} - \zeta q_{ii}^{-1}(d_i + \sum_{j<i} q_{ij}z_j^{(k+1)} + \sum_{j\geq i} q_{ij}z_j^{(k)}), u_i\}, \tag{2.13}$$

where mid$\{\cdot\}$ denotes the median of three numbers. When we apply (2.13), we cannot exploit the information of other variables, and hence the sequence may converge slowly. On the other hand, since the computation at each iteration is cheap, the total computational time may not be so long.

The sequence generated by the matrix splitting method for the linear complementarity problem is shown to converge globally under the following assumptions [3].

Assumption 1 Let (B, C) be a splitting of the matrix M.

(a) The matrix M is symmetric.

(b) The matrix $B - C$ is positive definite.

(c) For any k, a subproblem MCP$(B, C\boldsymbol{z}^{(k)} + \boldsymbol{d}, \boldsymbol{l}, \boldsymbol{u})$ has a solution.

(d) The following inequality holds:

$$\inf\left\{\frac{1}{2}\boldsymbol{z}^\top M\boldsymbol{z} + \boldsymbol{d}^\top\boldsymbol{z} \;\middle|\; \boldsymbol{l} \leq \boldsymbol{z} \leq \boldsymbol{u}\right\} > -\infty. \tag{2.14}$$

Assumption 1 holds for the SOR (2.13), when the following assumption holds.

Assumption 1'

(a) Assumption 1 (a) and (d) hold.

(b) The matrix M is positive semi-definite, all of the diagonal elements of M are positive and $0 < \zeta < 2$.

(c) The matrix $B - C$ is positive definite.

Under Assumption 1, we can easily show that the global convergence of the matrix splitting method for MCP$(F, \boldsymbol{l}, \boldsymbol{u})$. We omit the proof.

Theorem 1 Suppose that Assumption 1 holds. Then, every accumulation point of a sequence $\{z^{(k)}\}$ generated by the matrix splitting method is a solution of MCP$(F, \boldsymbol{l}, \boldsymbol{u})$.

In general, the matrix splitting method is poorly-convergent, hence it is not appropriate to obtaining an accurate solution. However, when we do not need the accurate solution, we can expect that the method has sufficient performance. Moreover, since the subproblem of the Jacobi method (2.10) is decomposed into small problem, it can be solved quickly by parallel computers.

3 Optimization Models for a Portfolio Function

In this section, we propose optimization models for constructing the portfolio function.

Suppose that an economical index $\boldsymbol{X_i}$ for the asset S_i are given. The example of $\boldsymbol{X_i}$ includes the prices and trading volume of the asset S_i, exchange rates of the currencies and some industry indices. We note that $\boldsymbol{X_i}$ is a data at the investment time while the return rate R_i is in the future. We treat both variables as random variables. Moreover, let $\boldsymbol{X} = (\boldsymbol{X_1}^\top, \ldots, \boldsymbol{X_n}^\top)^\top, R = (R_1, \ldots, R_n)^\top$, and let $\Omega_{\boldsymbol{X_i}}$ and $\Omega_{\boldsymbol{X}}$ be the sets of $\boldsymbol{X_i}$ and $\boldsymbol{X}$, respectively.

Let $g_i : \Omega_{\boldsymbol{X_i}} \to [0, 1]$ be a function from an index $\boldsymbol{X_i}$ to output the allocation of asset S_i, and define $\boldsymbol{g}(\boldsymbol{X}) = (g_1(\boldsymbol{X_1}), \ldots, g_n(\boldsymbol{X_n}))$. Then the portfolio $\boldsymbol{y}$ is given by

$$\boldsymbol{y} = (y_1, \ldots, y_n)^\top = \Big(g_1(\boldsymbol{X_1}), \ldots, g_n(\boldsymbol{X_n})\Big)^\top. \tag{3.1}$$

In this paper, we call the function $\boldsymbol{g}$ a portfolio function.

Now we formulate a minimization problem of LSAD with respect to the portfolio function $\boldsymbol{g}$. Substituting portfolio function (3.1) into the M-LSAD model (2.3), we obtain the following problem:

$$\begin{aligned} \text{minimize} \quad & E\Big[\max\{0, \alpha - \sum_{i=1}^{n} R_u g_i(\boldsymbol{X_i})\}\Big] \\ \text{subject to} \quad & g_i(\boldsymbol{X_i}) \geq 0, \quad i = 1, \ldots, n, \forall \boldsymbol{X} \subseteq \Omega_{\boldsymbol{X}}, \\ & \sum_{i=1}^{n} g_i(\boldsymbol{X_i}) \leq 1, \forall \boldsymbol{X} \subseteq \Omega_{\boldsymbol{X}}, \\ & E\Big[\sum_{i=1}^{n} R_i g_i(\boldsymbol{X_i})\Big] \geq \gamma. \end{aligned} \tag{3.2}$$

Note that the decision variable of the problem (3.2) is the function $\boldsymbol{g}$. Because of the following reasons, the problem (3.2) is difficult to solve.

- The probability distribution of $(\boldsymbol{X}, \boldsymbol{R})$ is unknown in general.

- Since the decision variable $\boldsymbol{g}$ is a function and the number of constraints are infinite, the problem (3.2) becomes the infinite programming problem.

To overcome the above difficulties, we propose an approximate model of the problem (3.2). In a way similar to deriving the M-LSAD model (2.3), we first approximate the expectations of random variables by using the historical data. Let economical indices $\boldsymbol{x_i^t}$ at the period $t = 1, 2, \ldots, T$ are given. Then we obtain the following approximation of the LSAD.

$$E\Big[\max\{0, \alpha - \sum_{i=1}^{n} R_i g_i(\boldsymbol{X_i})\}\Big] \approx \frac{1}{T}\max\{\alpha^t - \sum_{i=1}^{n} r_i^t g_i(\boldsymbol{x_i^t})\},$$

where α^t and r_i^t are the realizations given in Section 2. Similarly, the constraints of the problem (3.2) is represented as follows:

$$g_i(\boldsymbol{x_i^t}) \geq 0, \quad i = 1, 2, \ldots, n, \ t = 1, 2, \ldots, T,$$

$$\sum_{i=1}^{n} g_i(\boldsymbol{x_i^t}) \leq 1, \quad t = 1, 2, \ldots, T,$$

$$\frac{1}{T}\sum_{t=1}^{T}\sum_{i=1}^{n} r_i^t g_i(\boldsymbol{x_i^t}) \geq \gamma.$$

Consequently, the problem (3.2) is reduced to

$$\begin{aligned} \text{minimize} \quad & \sum_{t=1}^{T} \max\{0, \alpha^t - \sum_{i=1}^{n} r_i^t g_i(\boldsymbol{x_i^t})\} \\ \text{subject to} \quad & g_i(\boldsymbol{x_i^t}) \geq 0, \quad i = 1, 2 \ldots, n, \ t = 1, 2, \ldots, T, \\ & \sum_{i=1}^{n} g_i(\boldsymbol{x_i^t}) \leq 1, \quad t = 1, 2, \ldots, T, \\ & \frac{1}{T}\sum_{t=1}^{T}\sum_{i=1}^{n} g_i(\boldsymbol{x_i^t}) \geq \gamma. \end{aligned} \tag{3.3}$$

Note that the number of the constraints of this problem is finite. However, the decision variable is still the function $\boldsymbol{g}$, and hence it is still the semi-infinite programming problem. Thus we consider a restriction on the function $\boldsymbol{g}$ in order to represent $\boldsymbol{g}$ by the finite number of variable. Let $\boldsymbol{\phi_i}$ be a nonlinear mapping from the space $\Omega_{\boldsymbol{X_i}}$ to a feature space R^{θ_i}. We may allow $\theta_i = \infty$. The vector $\boldsymbol{\phi_i}$ is called a feature vector. We assume that for given $\boldsymbol{w_i} \in R^{\theta_i}$, the portfolio function $g_i(\boldsymbol{X_i})$ is represented as

$$g_i(\boldsymbol{X_i}) = \langle \boldsymbol{w_i}, \boldsymbol{\phi_i}(\boldsymbol{X_i}) \rangle, \tag{3.4}$$

where $\langle \cdot, \cdot \rangle$ denotes an appropriate inner product.

Remark 1 We can assume that the portfolio function is given by $g_i(\boldsymbol{X_i}) = \langle \boldsymbol{w_i}, \boldsymbol{\phi_i}(\boldsymbol{X_i}) \rangle + b_i$ with $b_i \in R$. The following discussion can be easily extended into the case.

By substituting this restriction (3.4) into the semi-infinite problem (3.3), we have the following problem:

$$\begin{aligned}
\text{minimize} \quad & \sum_{t=1}^{T} \max\{0, \alpha^t - \sum_{i=1}^{n} r_i^t \langle \boldsymbol{w_i}, \boldsymbol{\phi_i}(\boldsymbol{x_i^t}) \rangle\} \\
\text{subject to} \quad & \langle \boldsymbol{w_i}, \boldsymbol{\phi_i}(\boldsymbol{x_i^t}) \rangle \geq 0, \quad i = 1, 2, \ldots, n, \ t = 1, 2, \ldots, T, \\
& \sum_{i=1}^{n} \langle \boldsymbol{w_i}, \boldsymbol{\phi_i}(\boldsymbol{x_i^t}) \rangle \leq 1, \quad t = 1, 2, \ldots, T, \\
& \frac{1}{T} \sum_{t=1}^{T} \sum_{i=1}^{n} r_i^t \langle \boldsymbol{w_i}, \boldsymbol{\phi_i}(\boldsymbol{x_i^t}) \rangle \geq \gamma,
\end{aligned}$$

where the decision variable is $\boldsymbol{w} = (\boldsymbol{w_1}^\top, \ldots, \boldsymbol{w_n}^\top)^\top$. By introducing an auxiliary variable $\boldsymbol{\eta} = (\eta^1, \ldots, \eta^T)^\top$, it is equivalent to the following linear programming problem:

$$\begin{aligned}
\text{minimize} \quad & \frac{1}{T} \sum_{t=1}^{T} \eta^t \\
\text{subject to} \quad & \eta^t \geq \alpha^t - \sum_{i=1}^{n} r_i^t \langle \boldsymbol{w_i}, \boldsymbol{\phi_i}(\boldsymbol{x_i^t}) \rangle, \quad t = 1, 2, \ldots, T, \\
& \eta^t \geq 0, \quad t = 1, 2, \ldots, T, \\
& \langle \boldsymbol{w_i}, \boldsymbol{\phi_i}(\boldsymbol{x_i^t}) \rangle \geq 0, \ t = 1, \ldots, T, \ i = 1, 2, \ldots, n, \\
& \sum_{i=1}^{n} \langle \boldsymbol{w_i}, \boldsymbol{\phi_i}(\boldsymbol{x_i^t}) \rangle \leq 1, \quad t = 1, 2, \ldots, T, \\
& \frac{1}{T} \sum_{t=1}^{T} \sum_{i=1}^{n} r_i^t \langle \boldsymbol{w_i}, \boldsymbol{\phi_i}(\boldsymbol{x_i^t}) \rangle \geq \gamma.
\end{aligned} \tag{3.5}$$

The problem (3.5) is regarded as a learning problem of the function $\boldsymbol{g}$ when the historical data is given. If the dimension θ_i of the feature vector $\boldsymbol{\phi_i}$ is larger than the number of the data, then $\boldsymbol{\eta}$ can equal 0. It means that the return rate $\sum_{i=1}^{n} r_i^t g_i(\boldsymbol{x_i^t})$ is larger than the goal α^t for all periods. This phenomenon is called an overfitting in the field of machine learning. If the overfitting occurs, then the obtained result would be valid only for the given data, but not appropriate to unknown situations in the future. Namely, the result of assets management in the future can fall behind. One of the often-used techniques for preventing overfitting is the regularization [2]. The technique prevents the overfitting by adding a regularization term to the objective function. The various regularization terms are proposed. In what follows, we use $\tau \sum_{i=1}^{n} \|\boldsymbol{w}_i\|^2 / n$ as the term, where τ, called a regularization parameter, is a positive parameter for controlling the overfitting. Putting the term in the problem (3.5), we have the following

convex QP:

$$\begin{aligned}
\text{minimize} \quad & \frac{\tau}{n}\sum_{i=1}^{n}\|\boldsymbol{w_i}\|^2 + \frac{1}{T}\sum_{t=1}^{T}\eta^t \\
\text{subject to} \quad & \eta^t \geq \alpha^t - \sum_{i=1}^{n} r_i^t \langle \boldsymbol{w_i}, \boldsymbol{\phi_i}(\boldsymbol{x_i^t})\rangle, \quad t = 1, 2, \ldots, T, \\
& \eta^t \geq 0, \quad t = 1, 2, \ldots, T, \\
& \langle \boldsymbol{w_i}, \boldsymbol{\phi_i}(\boldsymbol{x_i^t})\rangle \geq 0, \ t = 1, \ldots, T, \ i = 1, 2, \ldots, n, \\
& \sum_{i=1}^{n}\langle \boldsymbol{w_i}, \boldsymbol{\phi_i}(\boldsymbol{x_i^t})\rangle \leq 1, \quad t = 1, 2, \ldots, T, \\
& \frac{1}{T}\sum_{t=1}^{T}\sum_{i=1}^{n} r_i^t \langle \boldsymbol{w_i}, \boldsymbol{\phi_i}(\boldsymbol{x_i^t})\rangle \geq \gamma.
\end{aligned} \tag{3.6}$$

Let $\boldsymbol{w^*}$ be the solution of the problem (3.6). Then, the portfolio function $g_i(\boldsymbol{X_i})$ is represented as follows.

$$g_i(\boldsymbol{X_i}) = \langle \boldsymbol{w_i^*}, \boldsymbol{\phi_i}(\boldsymbol{X_i})\rangle.$$

Remark 2 The constraints of the problem (3.6) hold for the historical data used during period $t = 1, 2, \ldots, T$. However, when a index $\boldsymbol{X} \subseteq \Omega_{\boldsymbol{X}}$ is different from any of the historical data $\boldsymbol{X^t}$, the constraints $g_i(\boldsymbol{X_i}) \geq 0$ and $\sum_{i=1}^{n} g_i(\boldsymbol{X_i}) \leq 1$ may not be satisfied. Therefore we have to transform $g_i(\boldsymbol{X_i})$ for the practical use. For example, when $g_i(\boldsymbol{X_i}) < 0$, we set $g_i(\boldsymbol{X_i}) = 0$. When $\sum_{i=1}^{n} g_i(\boldsymbol{X_i}) > 1$, we use a normalization $y_i = g_i(\boldsymbol{X_i}) / \sum_{i=1}^{n} g_i(\boldsymbol{X_i})$.

We next formulate the optimization model with CVaR minimization. Substituting (3.1) into the problem (2.4), we have

$$\begin{aligned}
\text{minimize}_{v,\boldsymbol{g}} \ & v + \frac{1}{(1-\beta)} E\left[\max\{0, f(\boldsymbol{g}(\boldsymbol{X}), \boldsymbol{R}) - v\}\right] \\
\text{subject to} \quad & g_i(\boldsymbol{X_i}) \geq 0, \quad i = 1, 2, \ldots, n, \quad \forall \boldsymbol{X} \subset \Omega_{\boldsymbol{X}}, \\
& \sum_{i=1}^{n} g_i(\boldsymbol{X_i}) \leq 1, \quad \forall \boldsymbol{X} \subset \Omega_{\boldsymbol{X}}, \\
& E\left[\sum_{i=1}^{n} R_i g_i(\boldsymbol{X_i})\right] \geq \gamma.
\end{aligned} \tag{3.7}$$

This problem is similar to the infinite programming problem based on the LSAD (3.2), and hence it is hard to obtain the solution. Then, in a way similar to deriving the problem (3.6), by using the historical

data and restricting the form of the function $\boldsymbol{g}$, we can reformulate (3.7) as the following convex QP.

$$\begin{aligned}
&\text{minimize}_{v,\boldsymbol{w},\boldsymbol{\eta}} \ \frac{\tau}{n}\sum_{i=1}^{n}\|\boldsymbol{w_i}\|^2 + v + \frac{1}{(1-\beta)T}\sum_{t=1}^{T}\eta^t \\
&\text{subject to} \quad \eta^t \geq -\sum_{i=1}^{n} r_i^t\langle \boldsymbol{w_i}, \boldsymbol{\phi_i}(\boldsymbol{x_i^t})\rangle - v, \quad t = 1,2,\ldots,T, \\
&\qquad\qquad \eta^t \geq 0, \quad t = 1,2,\ldots,T, \\
&\qquad\qquad \langle \boldsymbol{w_i}, \boldsymbol{\phi_i}(\boldsymbol{x_i^t})\rangle \geq 0, i = 1,2,\ldots,n, t = 1,2,\ldots,T, \\
&\qquad\qquad \sum_{i=1}^{n}\langle \boldsymbol{w_i}, \boldsymbol{\phi_i}(\boldsymbol{x_i^t})\rangle \leq 1, \ t = 1,2,\ldots,T, \\
&\qquad\qquad \frac{1}{T}\sum_{t=1}^{T}\sum_{i=1}^{n} r_i^t\langle \boldsymbol{w_i}, \boldsymbol{\phi_i}(\boldsymbol{x_i^t})\rangle \geq \gamma,
\end{aligned} \tag{3.8}$$

where the first term of the objective function is the regularization term for preventing the overfitting.

4 Dual Problems of the Proposed Models and Its Solutions with Kernel Functions

To solve the problem (3.6) given in Section 3, we have to evaluate the feature vector $\boldsymbol{\phi_i}$ for all $\boldsymbol{x_i^t}$. If the dimension of the feature vector $\boldsymbol{\phi_i}$ is very high or infinite, the problem (3.6) becomes the large-scale or infinite optimization problem, and hence we cannot solve the problem easily. Therefore, in this section, we consider dual problems of the problems (3.6) and (3.8). We will show that the number of the decision variables of the dual problem is in proportion to $n \times T$. Moreover, We will also see that we need not to evaluate $\boldsymbol{\phi_i}(\boldsymbol{x_i^t})$ by using the kernel methods and that the portfolio function $g_i(\boldsymbol{X_i})$ is represented by the linear combinations of some kernel functions.

The Lagrangian function [1] of the problem (3.6) is given by

$$\begin{aligned}
L(\boldsymbol{w},\boldsymbol{\eta},\boldsymbol{\kappa},\boldsymbol{\lambda},\boldsymbol{\mu},\boldsymbol{\nu},\xi) \ = \ & \frac{\tau}{n}\sum_{i=1}^{n}\|\boldsymbol{w_i}\|^2 + \frac{1}{T}\sum_{t=1}^{T}\eta^t + \sum_{t=1}^{T}\kappa^t\left\{\alpha^t - \sum_{i=1}^{n} r_i^t\langle \boldsymbol{w_i}, \boldsymbol{\phi_i}(\boldsymbol{x_i^t})\rangle - \eta^t\right\} \\
& - \sum_{t=1}^{T}\lambda^t\eta^t - \sum_{t=1}^{T}\sum_{i=1}^{n}\mu_i^t(\langle \boldsymbol{w_i}, \boldsymbol{\phi_i}(\boldsymbol{x_i^t})\rangle) + \sum_{t=1}^{T}\nu^t\left(\sum_{i=1}^{n}\langle \boldsymbol{w_i}, \boldsymbol{\phi_i}(\boldsymbol{x_i^t})\rangle - 1\right) \\
& + \xi\left\{\gamma - \frac{1}{T}\sum_{t=1}^{T}\sum_{i=1}^{n} r_i^t\langle \boldsymbol{w_i}, \boldsymbol{\phi_i}(\boldsymbol{x_i^t})\rangle\right\},
\end{aligned}$$

where $\boldsymbol{\kappa} \in R^T, \boldsymbol{\lambda} \in R^T, \boldsymbol{\mu} \in R^{nT}, \nu \in R^T, \xi \in R$ are the Lagrange multipliers corresponding to the constraints of the problem (3.6), and $\boldsymbol{\mu} = (\boldsymbol{\mu_1}^\top, \ldots, \boldsymbol{\mu_n}^\top)^\top$ with $\boldsymbol{\mu_i} = (\mu_i^1, \ldots, \mu_i^T)^\top$.

Let a function $\omega : R^{\{(n+2)T+1\}} \mapsto [-\infty, +\infty)$ be defined by

$$\omega(\boldsymbol{\kappa},\boldsymbol{\lambda},\boldsymbol{\mu},\boldsymbol{\nu},\xi) = \inf\{L(\boldsymbol{w},\boldsymbol{\eta},\boldsymbol{\kappa},\boldsymbol{\lambda},\boldsymbol{\mu},\boldsymbol{\nu},\xi) \mid (\boldsymbol{w}^\top, \boldsymbol{\eta}^\top)^\top \in R^{n+T}\}. \tag{4.1}$$

Then the Lagrangian dual problem of the problem (3.6) is given by

$$\begin{aligned}
&\text{maximize} \quad \omega(\boldsymbol{\kappa},\boldsymbol{\lambda},\boldsymbol{\mu},\boldsymbol{\nu},\xi) \\
&\text{subject to} \quad \boldsymbol{\kappa} \geq 0, \boldsymbol{\lambda} \geq 0, \boldsymbol{\mu} \geq 0, \boldsymbol{\nu} \geq 0, \xi \geq 0.
\end{aligned} \tag{4.2}$$

We now give the explicit formulation of the objective function ω as follows.

We rewrite the Lagrangian function L as

$$
\begin{aligned}
&L(\boldsymbol{w}, \boldsymbol{\eta}, \boldsymbol{\kappa}, \boldsymbol{\lambda}, \boldsymbol{\mu}, \boldsymbol{\nu}, \xi) \\
&= \frac{\tau}{n} \sum_{i=1}^{n} \|\boldsymbol{w_i}\|^2 + \sum_{i=1}^{n} \left\langle \boldsymbol{w_i}, -\sum_{t=1}^{T} \kappa^t r_i^t \boldsymbol{\phi_i}(\boldsymbol{x_i^t}) - \sum_{t=1}^{T} \mu_i^t \boldsymbol{\phi_i}(\boldsymbol{x_i^t}) + \sum_{t=1}^{T} \nu^t \boldsymbol{\phi_i}(\boldsymbol{x_i^t}) - \frac{\xi}{T} \sum_{t=1}^{T} r_i^t \boldsymbol{\phi_i}(\boldsymbol{x_i^t}) \right\rangle \\
&\quad + \sum_{t=1}^{T} \eta^t (\frac{1}{T} - \kappa^t - \lambda^t) + \sum_{t=1}^{T} \alpha^t \kappa^t - \sum_{t=1}^{T} \nu^t + \gamma\xi.
\end{aligned}
$$

Since L is decomposed into the terms of $\boldsymbol{w}$ and $\boldsymbol{\eta}$, we can compute the objective function ω by minimizing L for $\boldsymbol{w}$ and $\boldsymbol{\eta}$ separately. We first minimize L with respect to $\boldsymbol{w}$. Since L is a convex quadratic function with respect to $\boldsymbol{w}$, the following equation holds at the minimum $\boldsymbol{w_i^*}$ of L.

$$\nabla_w L(\boldsymbol{w}, \boldsymbol{\eta}, \boldsymbol{\kappa}, \boldsymbol{\lambda}, \boldsymbol{\mu}, \boldsymbol{\nu}, \xi) = 0.$$

The gradients of L for $\boldsymbol{w_i}$ $(i = 1, 2, \ldots, n)$ is

$$
\begin{aligned}
&\nabla_{\boldsymbol{w_i}} L(\boldsymbol{w}, \boldsymbol{\eta}, \boldsymbol{\kappa}, \boldsymbol{\lambda}, \boldsymbol{\mu}, \boldsymbol{\nu}, \xi) \\
&\quad = \frac{2\tau}{n} \boldsymbol{w_i} - \sum_{t=1}^{T} \kappa^t r_i^t \boldsymbol{\phi_i}(\boldsymbol{x_i^t}) - \sum_{t=1}^{T} \mu_i^t \boldsymbol{\phi_i}(\boldsymbol{x_i^t}) + \sum_{t=1}^{T} \nu^t \boldsymbol{\phi_i}(\boldsymbol{x_i^t}) - \frac{\xi}{T} \sum_{t=1}^{T} r_i^t \boldsymbol{\phi_i}(\boldsymbol{x_i^t}), \quad i = 1, 2, \ldots, n.
\end{aligned}
$$

By the above two equations, we have the minimizer $\boldsymbol{w_i^*}$ of $L(\cdot, \boldsymbol{\eta}, \boldsymbol{\kappa}, \boldsymbol{\lambda}, \boldsymbol{\mu}, \boldsymbol{\nu}, \xi)$.

$$\boldsymbol{w_i^*} = \frac{n}{2\tau} \left(\sum_{t=1}^{T} \kappa^t r_i^t \boldsymbol{\phi_i}(\boldsymbol{x_i^t}) + \sum_{t=1}^{T} \mu_i^t \boldsymbol{\phi_i}(\boldsymbol{x_i^t}) - \sum_{t=1}^{T} \nu^t \boldsymbol{\phi_i}(\boldsymbol{x_i^t}) + \frac{\xi}{T} \sum_{t=1}^{T} r_i^t \boldsymbol{\phi_i}(\boldsymbol{x_i^t}) \right), \quad i = 1, 2, \ldots, n. \tag{4.3}$$

Next we minimize L with respect to $\boldsymbol{\eta}$. Since L is a linear function with respect to η^t, we have

$$\inf\{\eta^t (\frac{1}{T} - \kappa^t - \lambda^t) | \eta^t \in R\} = \begin{cases} 0 & \text{if } \kappa^t + \lambda^t = \frac{1}{T}, \\ -\infty & \text{otherwise.} \end{cases}$$

Thus, when $\kappa^t + \lambda^t \neq 1/T$ for some t, we have $\omega(\boldsymbol{\kappa}, \boldsymbol{\lambda}, \boldsymbol{\mu}, \boldsymbol{\nu}, \xi) = -\infty$. Since the dual problem is the maximization problem, we need not to consider the above case. Thus, we consider only the case where $\kappa^t + \lambda^t = 1/T$ $(t = 1, 2, \ldots, T)$, and hence we add constraints $\kappa^t + \lambda^t = 1/T$ $(t = 1, 2, \ldots, T)$ to the dual problem (4.2).

Substituting $\kappa^t + \lambda^t = 1/T$ and $\boldsymbol{w}^*$ into L, we have

$$
\begin{aligned}
\omega(\boldsymbol{\kappa}, \boldsymbol{\lambda}, \boldsymbol{\mu}, \boldsymbol{\nu}, \xi) &= L(\boldsymbol{w}^*, \boldsymbol{\eta}, \boldsymbol{\kappa}, \boldsymbol{\lambda}, \boldsymbol{\mu}, \boldsymbol{\nu}, \xi) \\
&= -\frac{\tau}{n} \sum_{i=1}^{n} \|\boldsymbol{w_i^*}\|^2 + \sum_{t=1}^{T} \alpha^t \kappa^t - \sum_{t=1}^{T} \nu^t + \gamma\xi.
\end{aligned} \tag{4.4}
$$

Note that $\sum_{i=1}^{n} \|\boldsymbol{w_i^*}\|^2$ is written as

$$\sum_{i=1}^{n} \|\boldsymbol{w_i^*}\|^2 = \frac{n}{2\tau} \begin{pmatrix} \boldsymbol{\mu_1} \\ \vdots \\ \boldsymbol{\mu_n} \\ \boldsymbol{\kappa} \\ \boldsymbol{\nu} \\ \xi \end{pmatrix}^{\top} \bar{Q} \begin{pmatrix} \boldsymbol{\mu_1} \\ \vdots \\ \boldsymbol{\mu_n} \\ \boldsymbol{\kappa} \\ \boldsymbol{\nu} \\ \xi \end{pmatrix}, \tag{4.5}$$

where $\bar{Q} \in R^{\{(n+2)T+1\}\times\{(n+2)T+1\}}$ is the positive semi-definite matrix given by

$$\bar{Q} = \begin{pmatrix} H_1 & 0 & \cdots & 0 & H_1R_1 & -H_1 & \boldsymbol{p_1} \\ 0 & H_2 & \ddots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \ddots & \ddots & 0 & \vdots & \vdots & \vdots \\ 0 & \cdots & 0 & H_n & H_nR_n & -H_n & \boldsymbol{p_n} \\ R_1H_1 & \cdots & \cdots & R_nH_n & H_0 & -\sum_{i=1}^n R_iH_i & \boldsymbol{h} \\ -H_1 & \cdots & \cdots & -H_n & -\sum_{i=1}^n H_iR_i & \sum_{i=1}^n H_i & -\sum_{i=1}^n \boldsymbol{p_i} \\ \boldsymbol{p_1}^\top & \cdots & \cdots & \boldsymbol{p_n}^\top & \boldsymbol{h}^\top & -\sum_{i=1}^n \boldsymbol{p_i} & q \end{pmatrix}. \tag{4.6}$$

Here, R_i denotes the $T \times T$ diagonal matrix whose (t,t)-element is r_i^t, and $H_i,\ H_0 \in R^{T\times T},\ \boldsymbol{h},\ \boldsymbol{p_i} \in R^T,\ q \in R$ denote

$$\begin{aligned} &(H_i)_{kl} = \frac{n\langle \boldsymbol{\phi_i}(\boldsymbol{x_i^k}), \boldsymbol{\phi_i}(\boldsymbol{x_i^l})\rangle}{2\tau}, \quad H_0 = \sum_{i=1}^n R_iH_iR_i, \\ &\boldsymbol{h} = \frac{1}{T}\sum_{i=1}^n R_iH_i\boldsymbol{r_i}, \quad \boldsymbol{p_i} = \frac{1}{T}H_i\boldsymbol{r_i}, \quad q = \frac{1}{T^2}\sum_{i=1}^n \boldsymbol{r_i}^\top H_i\boldsymbol{r_i}, \\ &\boldsymbol{r_i} = (r_1^i, \ldots, r_i^T)^\top. \end{aligned}$$

From the above discussion, the dual problem (4.2) is written as

$$\begin{aligned} \text{maximize} \quad & -\frac{1}{2}\begin{pmatrix} \boldsymbol{\mu_1} \\ \vdots \\ \boldsymbol{\mu_n} \\ \boldsymbol{\kappa} \\ \boldsymbol{\nu} \\ \xi \end{pmatrix}^\top \bar{Q} \begin{pmatrix} \boldsymbol{\mu_1} \\ \vdots \\ \boldsymbol{\mu_n} \\ \boldsymbol{\kappa} \\ \boldsymbol{\nu} \\ \xi \end{pmatrix} + \sum_{t=1}^T \alpha^t\kappa^t - \sum_{t=1}^T \nu^t + \gamma\xi \\ \text{subject to} \quad & \kappa^t \geq 0, \lambda^t \geq 0, \kappa^t + \lambda^t = \frac{1}{T}, \quad t = 1, 2, \ldots, T, \\ & \boldsymbol{\mu} \geq 0,\ \boldsymbol{\nu} \geq 0,\ \xi \geq 0, \end{aligned}$$

which is equivalent to the convex QP:

$$\begin{aligned} \text{minimize} \quad & \frac{1}{2}\begin{pmatrix} \boldsymbol{\mu_1} \\ \vdots \\ \boldsymbol{\mu_n} \\ \boldsymbol{\kappa} \\ \boldsymbol{\nu} \\ \xi \end{pmatrix}^\top \bar{Q} \begin{pmatrix} \boldsymbol{\mu_1} \\ \vdots \\ \boldsymbol{\mu_n} \\ \boldsymbol{\kappa} \\ \boldsymbol{\nu} \\ \xi \end{pmatrix} - \sum_{t=1}^T \alpha^t\kappa^t + \sum_{t=1}^T \nu^t - \gamma\xi \\ \text{subject to} \quad & 0 \leq \kappa^t \leq \frac{1}{T}, \quad t = 1, 2, \ldots, T, \\ & \boldsymbol{\mu} \geq 0,\ \boldsymbol{\nu} \geq 0,\ \boldsymbol{\xi} \geq 0. \end{aligned} \tag{4.7}$$

We can express the optimal portfolio function $\boldsymbol{g}$ with the solution $(\hat{\boldsymbol{\mu}}, \hat{\boldsymbol{\kappa}}, \hat{\boldsymbol{\nu}}, \hat{\xi})$ of the dual problem (4.7). From the equation (4.3), we have

$$\boldsymbol{w_i^*} = \frac{n}{2\tau}\left(\sum_{t=1}^T \hat{\mu}_i^t\boldsymbol{\phi_i}(\boldsymbol{x_i^t}) + \sum_{t=1}^T \hat{\kappa}^t r_i^t\boldsymbol{\phi_i}(\boldsymbol{x_i^t}) - \sum_{t=1}^T \hat{\nu}^t\boldsymbol{\phi_i}(\boldsymbol{x_i^t}) + \frac{\hat{\xi}}{T}\sum_{t=1}^T r_i^t\boldsymbol{\phi_i}(\boldsymbol{x_i^t})\right), \quad i = 1, 2, \ldots, n.$$

Hence we can represent the optimal portfolio function (3.4) as follows:

$$\begin{aligned} g_i(\boldsymbol{X_i}) &= \langle \boldsymbol{w_i^*}, \boldsymbol{\phi_i}(\boldsymbol{X_i}) \rangle \\ &= \frac{n}{2\tau} \left(\sum_{t=1}^{T} \hat{\mu}_i^t \langle \boldsymbol{\phi_i}(\boldsymbol{x_i^t}), \boldsymbol{\phi_i}(\boldsymbol{X_i}) \rangle + \sum_{t=1}^{T} \hat{\kappa}^t r_i^t \langle \boldsymbol{\phi_i}(\boldsymbol{x_i^t}), \boldsymbol{\phi_i}(\boldsymbol{X_i}) \rangle \right. \\ &\quad \left. - \sum_{t=1}^{T} \hat{\nu}^t \langle \boldsymbol{\phi_i}(\boldsymbol{x_i^t}), \boldsymbol{\phi_i}(\boldsymbol{X_i}) \rangle + \frac{\hat{\xi}}{T} \sum_{t=1}^{T} r_i^t \langle \boldsymbol{\phi_i}(\boldsymbol{x_i^t}), \boldsymbol{\phi_i}(\boldsymbol{X_i}) \rangle \right), \quad i = 1, 2, \ldots, n. \end{aligned} \tag{4.8}$$

Remark 3 As discussed in Remark 2 in Section 3, the constraints of problem (3.3) are satisfied only for the historical data during period $t = 1, \ldots, T$. Thus, the constraints $g_i(\boldsymbol{X_i}) \geq 0$, $\sum_{i=1}^{n} g_i(\boldsymbol{X_i}) \leq 1$ may not be satisfied for the current economical index $\boldsymbol{X}$. If the constraints is not satisfied, then we need to transform $\boldsymbol{g}$ as discussed in Remark 2.

The feature vector $\boldsymbol{\phi_i}(\boldsymbol{X_i})$ does not appear solely in the coefficient matrix $\bar{Q}$ of the dual problem (4.7) and in the optimal portfolio function $\boldsymbol{g}$ given by (4.8). On the other hand, the inner products of $\boldsymbol{\phi_i}(\boldsymbol{X_i})$ and $\boldsymbol{\phi_i}(\boldsymbol{x_i^t})$ appear. Now, let $K_i : \Omega_{\boldsymbol{X_i}} \times \Omega_{\boldsymbol{X_i}} \to R.$ be given by

$$K_i(\boldsymbol{x_i}, \bar{\boldsymbol{x}}_{\boldsymbol{i}}) = \langle \boldsymbol{\phi_i}(\boldsymbol{x_i}), \boldsymbol{\phi_i}(\bar{\boldsymbol{x}}_{\boldsymbol{i}}) \rangle.$$

Then, the matrix $\bar{Q}$ is represented with the function K_i only. Moreover, the optimal portfolio function $\boldsymbol{g}$ is also given with K_i.

$$g_i(\boldsymbol{X_i}) = \frac{n}{2\tau} \left(\sum_{t=1}^{T} \hat{\mu}_i^t K_i(\boldsymbol{x_i^t}, \boldsymbol{X_i}) + \sum_{t=1}^{T} \hat{\kappa}^t r_i^t K_i(\boldsymbol{x_i^t}, \boldsymbol{X_i}) - \sum_{t=1}^{T} \hat{\nu}^t K_i(\boldsymbol{x_i^t}, \boldsymbol{X_i}) + \frac{\hat{\xi}}{T} \sum_{t=1}^{T} r_i^t K_i(\boldsymbol{x_i^t}, \boldsymbol{X_i}) \right). \tag{4.9}$$

Conversely, if we first define a function $K_i : \Omega_{\boldsymbol{X_i}} \times \Omega_{\boldsymbol{X_i}} \to R$, then we can construct the matrix $\bar{Q}$ in the dual problem (4.7) without defining the feature vector $\boldsymbol{\phi_i}(\boldsymbol{X_i})$ explicitly. However, the dual problem (4.7) with K_i may not be the convex QP. Moreover, $\boldsymbol{g}$ given in the form of (4.9) with K_i may not be a solution of the original problem (3.6). A sufficient condition for the convexity and the equivalence is that the function K_i is a kernel function, which is defined as follows [11].

Definition 1 A function $K : X \times X \to R$ is said to be a kernel function, if there exists a feature vector $\boldsymbol{\phi} : X \mapsto F$, and the following equation is satisfied for all $\boldsymbol{x}, \bar{\boldsymbol{x}} \in X$.

$$K(\boldsymbol{x}, \bar{\boldsymbol{x}}) = \langle \boldsymbol{\phi}(\boldsymbol{x}), \boldsymbol{\phi}(\bar{\boldsymbol{x}}) \rangle.$$

The technique obtaining the optimal solution with kernel functions is said to be the kernel trick [11]. The kernel trick is one of the main techniques in the kernel methods.

The following are popular examples of the kernel function.

$$K(\boldsymbol{x}, \bar{\boldsymbol{x}}) = (\boldsymbol{x}^\top \bar{\boldsymbol{x}} + c)^d \quad (d : \text{natural number}, \ c \geq 0). \tag{4.10}$$

$$K(\boldsymbol{x}, \bar{\boldsymbol{x}}) = \exp\left(- \frac{\|\boldsymbol{x} - \bar{\boldsymbol{x}}\|^2}{\sigma^2}. \right) \tag{4.11}$$

The function (4.10) is called a polynomial kernel, and (4.11) is called a Gaussian kernel. The dimension of the feature vector for the polynomial kernel is finite. On the other hand, the dimension of the feature vector for the Gaussian kernel is infinite.

We next give the dual problem of the CVaR minimization model (3.8).

The Lagrangian function of problem (3.8) is given by

$$
\begin{aligned}
&L(\boldsymbol{w}, \boldsymbol{\eta}, \boldsymbol{\kappa}, \boldsymbol{\lambda}, \boldsymbol{\mu}, \boldsymbol{\nu}, \xi) \\
&= \frac{\tau}{n}\sum_{i=1}^{n}\|\boldsymbol{w_i}\|^2 + v + \frac{1}{(1-\beta)T}\sum_{t=1}^{T}\eta^t + \sum_{t=1}^{T}\kappa^t\Big\{-\sum_{i=1}^{n} r_i^t\langle \boldsymbol{w_i}, \boldsymbol{\phi_i}(\boldsymbol{x_i^t})\rangle - v - \eta^t\Big\} \\
&\quad -\sum_{t=1}^{T}\lambda^t\eta^t - \sum_{t=1}^{T}\sum_{i=1}^{n}\mu_i^t(\langle \boldsymbol{w_i}, \boldsymbol{\phi_i}(\boldsymbol{x_i^t})\rangle) + \sum_{t=1}^{T}\nu^t\Big(\sum_{i=1}^{n}\langle \boldsymbol{w_i}, \boldsymbol{\phi_i}(\boldsymbol{x_i^t})\rangle - 1\Big) \\
&\quad +\xi\Big\{\gamma - \frac{1}{T}\sum_{t=1}^{T}\sum_{i=1}^{n} r_i^t\langle \boldsymbol{w_i}, \boldsymbol{\phi_i}(\boldsymbol{x_i^t})\rangle\Big\} \\
&= \frac{\tau}{n}\sum_{i=1}^{n}\|\boldsymbol{w_i}\|^2 + \sum_{i=1}^{n}\Big\langle \boldsymbol{w_i}, -\sum_{t=1}^{T}\kappa^t r_i^t\boldsymbol{\phi_i}(\boldsymbol{x_i^t}) - \sum_{t=1}^{T}\mu_i^t\boldsymbol{\phi_i}(\boldsymbol{x_i^t}) + \sum_{t=1}^{T}\nu^t\boldsymbol{\phi_i}(\boldsymbol{x_i^t}) - \frac{\xi}{T}\sum_{t=1}^{T} r_i^t\boldsymbol{\phi_i}(\boldsymbol{x_i^t})\Big\rangle \\
&\quad +\sum_{t=1}^{T}\eta^t\left(\frac{1}{(1-\beta)T} - \kappa^t - \lambda^t\right) + \left(1 - \sum_{t=1}^{T}\kappa^t\right)v - \sum_{t=1}^{T}\nu^t + \gamma\xi, \qquad (4.12)
\end{aligned}
$$

where $\boldsymbol{\kappa}, \boldsymbol{\lambda}, \boldsymbol{\mu}, \boldsymbol{\nu}$ and ξ are the Lagrange multipliers corresponding to the constraints of the problem (3.8). The Lagrangian dual problem of the problem (3.8) is given by (4.2).

The Lagrangian function L is decomposed into functions of $\boldsymbol{w}, \boldsymbol{\eta}$ and v, respectively. Hence, minimizing L with respect to $\boldsymbol{w}, \boldsymbol{\eta}, v$ separately, we can obtain the objective function value of the dual problem. We first minimize L with respect to $\boldsymbol{w}$. The gradients of L for w_i $(i = 1, 2, \ldots, n)$ is

$$
\begin{aligned}
&\nabla_{\boldsymbol{w_i}} L(\boldsymbol{w}, \boldsymbol{\eta}, \boldsymbol{\kappa}, \boldsymbol{\lambda}, \boldsymbol{\mu}, \boldsymbol{\nu}, \xi) \\
&\quad = \frac{2\tau}{n}\boldsymbol{w_i} - \sum_{t=1}^{T}\kappa^t r_i^t\boldsymbol{\phi_i}(\boldsymbol{x_i^t}) - \sum_{t=1}^{T}\mu_i^t\boldsymbol{\phi_i}(\boldsymbol{x_i^t}) + \sum_{t=1}^{T}\nu^t\boldsymbol{\phi_i}(\boldsymbol{x_i^t}) - \frac{\xi}{T}\sum_{t=1}^{T} r_i^t\boldsymbol{\phi_i}(\boldsymbol{x_i^t}), \quad i = 1, 2, \ldots, n.
\end{aligned}
$$

Thus, we have the minimizer $\boldsymbol{w_i^*}$ of $L(\cdot, \boldsymbol{\eta}, \boldsymbol{\kappa}, \boldsymbol{\lambda}, \boldsymbol{\mu}, \boldsymbol{\nu}, \xi)$.

$$
\boldsymbol{w_i^*} = \frac{n}{2\tau}\Big(\sum_{t=1}^{T}\kappa^t r_i^t\boldsymbol{\phi_i}(\boldsymbol{x_i^t}) + \sum_{t=1}^{T}\mu_i^t\boldsymbol{\phi_i}(\boldsymbol{x_i^t}) - \sum_{t=1}^{T}\nu^t\boldsymbol{\phi_i}(\boldsymbol{x_i^t}) + \frac{\xi}{T}\sum_{t=1}^{T} r_i^t\boldsymbol{\phi_i}(\boldsymbol{x_i^t})\Big), \quad i = 1, \ldots, n.
$$

We next minimize L with respect to $\boldsymbol{\eta}$. Since L is the linear function with respect to $\boldsymbol{\eta}$, L tends to be $-\infty$ as $\eta^t \to +\infty$ when $\kappa^t + \lambda^t \neq 1/(1-\beta)T$ for some t. Thus, we consider only the case where $\kappa^t + \lambda^t = 1/(1-\beta)T$ for all t, and we add these equations to the dual problem (4.2) as the constraints.

Since L is also the linear function with respect to v, we add a constraint $\sum_{t=1}^{T}\kappa^t = 1$ to the dual problem (4.2).

Substituting $\kappa^t + \lambda^t = 1/(1-\beta)T$, $\sum_{t=1}^{T}\kappa^t = 1$ and $\boldsymbol{w}^*$ into L, we have

$$
\omega(\boldsymbol{\kappa}, \boldsymbol{\lambda}, \boldsymbol{\mu}, \boldsymbol{\nu}, \xi) = L(\boldsymbol{w}^*, \boldsymbol{\eta}, \boldsymbol{\kappa}, \boldsymbol{\lambda}, \boldsymbol{\mu}, \boldsymbol{\nu}, \xi) = -\frac{\tau}{n}\sum_{i=1}^{n}\|\boldsymbol{w_i^*}\|^2 - \sum_{t=1}^{T}\nu^t + \gamma\xi.
$$

Hence we can write the dual problem (4.2) is written as

$$\begin{aligned}
\text{minimize} \quad & \frac{1}{2}\begin{pmatrix} \boldsymbol{\mu_1} \\ \vdots \\ \boldsymbol{\mu_n} \\ \boldsymbol{\kappa} \\ \boldsymbol{\nu} \\ \xi \end{pmatrix}^{\top} \bar{Q} \begin{pmatrix} \boldsymbol{\mu_1} \\ \vdots \\ \boldsymbol{\mu_n} \\ \boldsymbol{\kappa} \\ \boldsymbol{\nu} \\ \xi \end{pmatrix} + \sum_{t=1}^{T} \nu^t - \gamma\xi \\
\text{subject to} \quad & 0 \leq \kappa^t \leq \frac{1}{(1-\beta)T}, \quad t = 1, 2, \ldots, T, \\
& \sum_{t=1}^{T} \kappa^t = 1, \\
& \boldsymbol{\mu} \geq 0, \ \boldsymbol{\nu} \geq 0, \ \xi \geq 0,
\end{aligned} \tag{4.13}$$

where the matrix $\bar{Q}$ is defined as (4.6). By using the solution $(\hat{\boldsymbol{\mu}}, \hat{\boldsymbol{\kappa}}, \hat{\boldsymbol{\nu}}, \hat{\xi})$ of the dual problem (4.13), we can represent the solution of the CVaR minimization model (3.8) as the equation (4.8). We can also use the kernel trick for the CVaR minimization model, because the matrix $\bar{Q}$ in the dual problem and the portfolio function $\boldsymbol{g}$ are represented with kernel functions.

5 The Matrix Splitting Method for the Dual Problem

When the numbers of the assets n and the period T are large, the problem (4.7) becomes very large-scale. Thus, we consider the matrix splitting method described in Section 2 to solve (4.7). Note that, since the data used are usually insufficient and the problems include some uncertainties, we need not to obtain an accurate solution of the problem (4.7). Moreover, as written below, the problem can be parallelized by exploiting the structure $\bar{Q}$ defined by (4.6). Therefore the matrix splitting method is appropriate to the problem (4.7). Now we discuss the implemental.

The KKT conditions of the problem (4.7) are written as

$$\begin{aligned}
& \bar{Q}\bar{\boldsymbol{z}} + \bar{\boldsymbol{d}} - \boldsymbol{\lambda_1} + \boldsymbol{\lambda_2} = 0, \ \boldsymbol{\lambda_1}^{\top}\bar{\boldsymbol{z}} = 0, \ \boldsymbol{\lambda_2}^{\top}(\bar{\boldsymbol{u}} - \bar{\boldsymbol{z}}) = 0, \\
& \bar{\boldsymbol{z}} \geq 0, \boldsymbol{\lambda_1} \geq 0, \\
& \bar{\boldsymbol{z}} \leq \bar{\boldsymbol{u}}, \boldsymbol{\lambda_2} \geq 0,
\end{aligned}$$

where

$$\begin{aligned}
& \bar{\boldsymbol{z}} = (\boldsymbol{\mu_1}^{\top}, \ldots, \boldsymbol{\mu_n}^{\top}, \boldsymbol{\kappa}^{\top}, \boldsymbol{\nu}^{\top}, \xi)^{\top}, \\
& \bar{\boldsymbol{d}} = (0, \ldots, 0, -\alpha^1, \ldots, -\alpha^T, 1, \ldots, 1, -\gamma)^{\top}, \\
& \bar{u}_i = \begin{cases} \frac{1}{T} & \text{if } i = nT+1, \ldots, (n+1)T, \\ +\infty & \text{otherwise.} \end{cases}
\end{aligned}$$

Then, these KKT conditions are equivalent to MCP$(\bar{Q}, \bar{d}, \mathbf{0}, \bar{\boldsymbol{u}})$, where $\mathbf{0} = (0, \ldots, 0)$. Let (B, C) be a

splitting of $\bar{Q}$ given by

$$B = \begin{pmatrix} \frac{1}{\zeta}H_1 & & & & & & \\ 0 & \frac{1}{\zeta}H_2 & & & & & \\ \vdots & \ddots & \ddots & & & & \\ 0 & \cdots & 0 & \frac{1}{\zeta}H_n & & & \\ R_1H_1 & \cdots & \cdots & R_nH_n & \frac{1}{\zeta}H_0 & & \\ -H_1 & \cdots & \cdots & -H_n & -\sum_{i=1}^n H_iR_i & \frac{1}{\zeta}\sum_{i=1}^n H_i & \\ \boldsymbol{p_1}^\top & \cdots & \cdots & \boldsymbol{p_n}^\top & \boldsymbol{h}^\top & -\sum_{i=1}^n \boldsymbol{p_i} & \frac{1}{\zeta}q \end{pmatrix},$$

$$C = \begin{pmatrix} (1-\frac{1}{\zeta})H_1 & 0 & \cdots & 0 & H_1R_1 & -H_1 & \boldsymbol{p_1} \\ & (1-\frac{1}{\zeta})H_2 & \ddots & \vdots & \vdots & \vdots & \vdots \\ & & \ddots & 0 & \vdots & \vdots & \vdots \\ & & & (1-\frac{1}{\zeta})H_n & H_nR_n & -H_n & \boldsymbol{p_n} \\ & & & & (1-\frac{1}{\zeta})H_0 & -\sum_{i=1}^n R_iH_i & \boldsymbol{h} \\ & & & & & (1-\frac{1}{\zeta})\sum_{i=1}^n H_i & -\sum_{i=1}^n \boldsymbol{p_i} \\ & & & & & & (1-\frac{1}{\zeta})q \end{pmatrix}. \tag{5.1}$$

By using (B, C), the block-SOR method solves the following n+3 MCPs at each iteration successively.

$$\begin{aligned}
&\mathrm{MCP}(\zeta^{-1}H_i,\ (1-\zeta^{-1})H_i\boldsymbol{\mu_i}^{(k)} + H_iR_i\boldsymbol{\kappa}^{(k)} - H_i\boldsymbol{\nu}^{(k)} + \xi^{(k)}\boldsymbol{p_i},\ \mathbf{0}, +\boldsymbol{\infty}),\quad i = 1, 2, \ldots, n,\\
&\mathrm{MCP}(\zeta^{-1}H_0,\ \sum_{i=1}^n R_iH_i\boldsymbol{\mu_i}^{(k+1)} + (1-\zeta^{-1})H_0\boldsymbol{\kappa}^{(k)} - \sum_{i=1}^n R_iH_i\boldsymbol{\nu}^{(k)} + \xi^{(k)}\boldsymbol{h} - \boldsymbol{\alpha},\ \mathbf{0},\ T^{-1}\boldsymbol{e}),\\
&\mathrm{MCP}(\zeta^{-1}\sum_{i=1}^n H_i,\ -\sum_{i=1}^n H_i\boldsymbol{\mu_i}^{(k+1)} - \sum_{i=1}^n H_iR_i\boldsymbol{\kappa}^{(k+1)} + (1-\zeta^{-1})\sum_{i=1}^n H_i\boldsymbol{\nu}^{(k)} - \xi^{(k)}\sum_{i=1}^n \boldsymbol{p_i} + \boldsymbol{e},\ \mathbf{0},\ +\boldsymbol{\infty}),\\
&\mathrm{MCP}(\zeta^{-1}q,\ \sum_{i=1}^n \boldsymbol{p_i}^\top\boldsymbol{\mu_i}^{(k+1)} + \boldsymbol{h}^\top\boldsymbol{\kappa}^{(k+1)} - \sum_{i=1}^n \boldsymbol{p_i}\boldsymbol{\nu}^{(k+1)} + (1-\zeta^{-1})\xi^{(k)} - \gamma,\ 0,\ \infty),
\end{aligned} \tag{5.2}$$

where $+\boldsymbol{\infty} = (+\infty, \ldots, +\infty)^\top$ and $\boldsymbol{e} = (1, \ldots, 1)^\top$. Note that the decision variables of these MCPs are $\boldsymbol{\mu_1}$, $\boldsymbol{\mu_2}$, $\ldots$, $\boldsymbol{\mu_n}$, $\boldsymbol{\kappa}$, $\boldsymbol{\nu}$ and ξ, respectively. Thus the maximum size of the MCPs is T. Furthermore, we can solves the first n subproblems separately. Hence, using parallel processing according to the number of the assets, the computational cost at each iteration is propose to T. Moreover, when we split $\bar{Q}$ in each variable, that is, we apply (2.13), $\bar{z}_i^{(k+1)}$ is obtained by the following equations:

$$\bar{z}_i^{(k+1)} = \begin{cases} \mathrm{mid}\{0, \bar{z}_i^{(k)} - \zeta\bar{q}_{ii}^{-1}(\bar{d}_i + \sum_{j<i}\bar{q}_{ij}\bar{z}_j^{(k+1)} + \sum_{j\geq i}\bar{q}_{ij}\bar{z}_j^{(k)}), \frac{1}{T}\} & \text{if } i = nT+1, \ldots, (n+1)T \\ \max\{0, \bar{z}_i^{(k)} - \zeta\bar{q}_{ii}^{-1}(\bar{d}_i + \sum_{j<i}\bar{q}_{ij}\bar{z}_j^{(k+1)} + \sum_{j\geq i}\bar{q}_{ij}\bar{z}_j^{(k)})\} & \text{otherwise,} \end{cases} \tag{5.3}$$

where $\bar{q}_{ij}$ denotes (i, j) element of the matrix $\bar{Q}$. As described in Section 2, though the number of iterations may be large, each computational cost of (5.3) is very cheap.

We now discuss the global convergence of the matrix splitting method (5.3). First the matrix $\bar{Q}$ is symmetric positive semi-definite, and the problem (4.7) has a solution, therefore Assumption 1' (a) holds. Moreover, all of the diagonal elements of the matrix $\bar{Q}$ are positive, if we use the positive definite

kernels such as the polynomial kernel (4.10) and the Gaussian kernel (4.11). Thus Assumption 1' (b) holds when we choose ζ such that $0 < \zeta < 2$. Finally, since $\boldsymbol{v}^\top(B-C)\boldsymbol{v} = \boldsymbol{v}^\top D\boldsymbol{v}$ for all $\boldsymbol{v}$, $B-C$ is positive definite. Therefore Assumption 1' (c) holds. Consequently, by Theorem 1, we can obtain the global optimal solution of the dual problem (4.7) by using the procedure (5.3).

Unfortunately, the same matrix splitting method for the CVaR minimization model (4.13) does not satisfy Assumption 1' because the problem (4.13) have the equality constraint. Then we may use the coordinate gradient descent method [13], which is one of the efficient solver for such a large-scale QP with one equality constraint.

6 Numerical Experiments

In this section, we report results of numerical experiments for the proposed models. We compared the proposed model (3.6) with the M-LSAD model (2.3). All computation were carried out in Matlab 7.4 on a machine with 3.2 GHz Pentium 4 CPU and 2.0 GB memory.

We chose 191 stocks included in NIKKEI 225 as the objective assets. All of the data of the stocks were collected from Yahoo! Finance [14], The economical indices $\boldsymbol{X}$ were calculated from trading volumes of some stocks as follows. We first chose some categories of industry, that is, the food, the electronics device, the car and the communications. Let C_k be a set of assets that belongs to a category k. Let v_i^t be the rate of the trading volume of the asset S_i during the period t. We define V_k^t, $k = 1, 2, 3, 4$ as the means of the rate of the assets in the categories, that is,

$$V_k^t = \frac{1}{|C_k|} \sum_{i \in C_k} v_i^t,$$

where $|C_k|$ denotes the number of elements in C_k. Moreover, let V^t be the mean of the trading volumes of all of the stocks during the period t, and we define $\tilde{V}_k^t$, $k = 1, 2, 3, 4$ as the difference between V_k^t and V^t. Then, $\tilde{V}_k^t$ is given by

$$\tilde{V}_k^t = V_k^t - V^t.$$

Furthermore, let I_k^t be the means of $\tilde{V}_k^t$ of the recent four weeks, that is,

$$I_k^t = \frac{1}{4}(\tilde{V}_k^t + \tilde{V}_k^{t-1} + \tilde{V}_k^{t-2} + \tilde{V}_k^{t-3}).$$

Then, we define the realization of the economical indices $\boldsymbol{x}^t$ during the period t as

$$\boldsymbol{x}^t = (I_1^t,\ I_2^t,\ I_3^t,\ I_4^t)^\top. \tag{6.1}$$

We adopted the SOR method (5.3) as the matrix splitting method for solving the MCP$(\bar{Q}, \bar{d}, \boldsymbol{0}, \bar{\boldsymbol{u}})$. We employed

$$\|\boldsymbol{\Phi}(\boldsymbol{z}^{(k)})\| < \sqrt{\varepsilon} \tag{6.2}$$

as the stopping rule, where ε is a stopping parameter and $\boldsymbol{\Phi}(\boldsymbol{z})$ is given by

$$\boldsymbol{\Phi}(\boldsymbol{z}) = (\Phi_1(\boldsymbol{z}), \ldots, \Phi_{(n+2)T+1}(\boldsymbol{z}))^\top,$$

$$\Phi_i(\boldsymbol{z}) = \mathrm{mid}\{z_i, z_i - \bar{u}_i, (\bar{Q}z + \bar{d})_i\}.$$

Note that $\|\cdot\|$ denotes the L_2 norm of a vector.

We set the regularization parameter $\tau = 0.05$. Moreover, we set the realization of the Nikkei Stock Average during period t ($t = 1, \ldots, T$) as α^t. We adopted the Gaussian kernel (4.11) as the kernel function, and we set the Gaussian parameter as $\sigma^2 = 50$.

We used two sets of the data, Data 1 and Data 2 as in Table 1. Data 1 includes as the training data from 2000 to 2004, and as the test data from 2005 to 2006. Data 2 includes as the training data from 2001 to 2005, and as the test data from 2006 to 2007. The return rate of 191 stocks were calculated from the weekly data.

Table 1 The data we use in the experiments

	The training data	The test data
Data 1	the weekly data during '00 - '04	the weekly data during '05 - '06
Data 2	the weekly data during '01 - '05	the weekly data during '06 - '07

6.1 Experiments with Various Parameters for the Matrix Splitting Method

In the first experiment, we investigated the effect of relaxation parameter ζ in the SOR method (2.13) for the proposed model. Table 2 reports the number of iterations and the calculation time in which we obtain a solution satisfying (6.2) with $\varepsilon = 1.0 \times 10^{-5}$. We used Data 1. From the results, we see that we can obtain the solution in the practical time, if we choose ζ appropriately. In particular, $\zeta = 1.2$ is the best among the experiments. Hence, we choose $\zeta = 1.2$ in the subsequent experiments.

Table 2 Effect of relaxation parameter

ζ	Iteration $\gamma = 0.005$	Time (s)	Iteration $\gamma = 0.010$	Time (s)
0.8	173	552.23	165	551.33
0.9	157	498.08	143	470.01
1.0	146	466.01	133	435.76
1.1	135	427.54	128	413.64
1.2	127	405.27	128	416.23
1.3	123	388.25	129	419.56
1.4	123	388.81	133	449.67
1.5	127	396.35	141	464.58
1.6	149	491.19	995	3440.3

We next investigated the effect of the stopping parameter ε. The results are shown in Tables 3 and 4. Note that 'Mean' and 'Std' in the tables denote the expectation and the standard deviation of the return rate per a week, respectively, and 'Sharpe' denotes Sharpe ratio given by Mean/Std. From these results, we see that the method with $\varepsilon = 1.0 \times 10^{-6}$ takes more iterations than that with $\varepsilon = 1.0 \times 10^{-5}$, while the performances of the portfolios obtained from $\varepsilon = 1.0 \times 10^{-5}$ and 1.0×10^{-6} are not much different. This indicates that we need not to employ the restrictive stopping rule for appropriate solutions. Therefore, we set the stopping parameter $\varepsilon = 1.0 \times 10^{-5}$ in the subsequent experiments.

Table 3 Effect of stopping parameter : Data 1

ε	Iteration	Mean ($\times 10^{-3}$)	Std ($\times 10^{-3}$)	Sharpe
1.0×10^{-3}	57	5.09	12.8	0.398
1.0×10^{-4}	90	5.17	12.9	0.402
1.0×10^{-5}	122	5.20	12.9	0.403
1.0×10^{-6}	166	5.19	12.9	0.403

Table 4 Effect of stopping parameter : Data 2

ε	Iteration	Mean ($\times 10^{-3}$)	Std ($\times 10^{-3}$)	Sharpe
1.0×10^{-3}	96	1.55	16.4	0.0944
1.0×10^{-4}	166	1.49	16.9	0.0878
1.0×10^{-5}	253	1.44	17.1	0.0844
1.0×10^{-6}	890	1.50	16.7	0.0895

6.2 Comparisons of the Proposed Model with the M-LSAD Model

We next compared the proposed model (3.6) with the M-LSAD model (2.3) for various values of the required return rate γ. Table 5 shows the results of the portfolios obtained from the proposed model and the M-LSAD model, respectively. 'Average share of the top stock' denotes the average share of the stock which has the highest average share at each period. 'Average share of the top 5 stocks' is the same average of top 5 share stock. 'Average share of stocks' denotes the average of the share of all stocks. From the table, we see that the value of the standard deviation with the proposed model was less than that with the M-LSAD model. The reason for the result is that the share of each stock bought by the proposed model was less, while the number of the bought stocks was much more than that in the M-LSAD model. Thus, the proposed model is more risk-adverse than the M-LSAD model. Moreover, the proposed

Table 5 Results of the portfolios ($\gamma = 0.005$)

Data set	Data 1		Data2	
Portfolio	proposed	M-LSAD	proposed	M-LSAD
Mean ($\times 10^{-3}$)	5.20	4.79	1.44	-0.167
Std ($\times 10^{-3}$)	12.9	22.1	17.1	21.5
Sharpe	0.403	0.227	0.0844	-0.00778
Average share of the top stock	0.0427	0.311	0.0450	0.292
Average share of the top 5 stocks	0.0950	0.770	0.102	0.663
Average share of stocks	0.600	0.844	0.631	0.970

model was superior to the M-LSAD model with respect to the value of the Sharpe ratio. Particularly, the mean of the return rate of the M-LSAD model was negative for the Data2. On the other hand, that of the proposed model was positive. The result shows that we were able to obtain a portfolio according to situations by the proposed model.

Figures 1 and 2 are show the fluctuation of the value of the portfolios with required return rate $\gamma = 0.005$ for Data1 and Data 2, respectively.

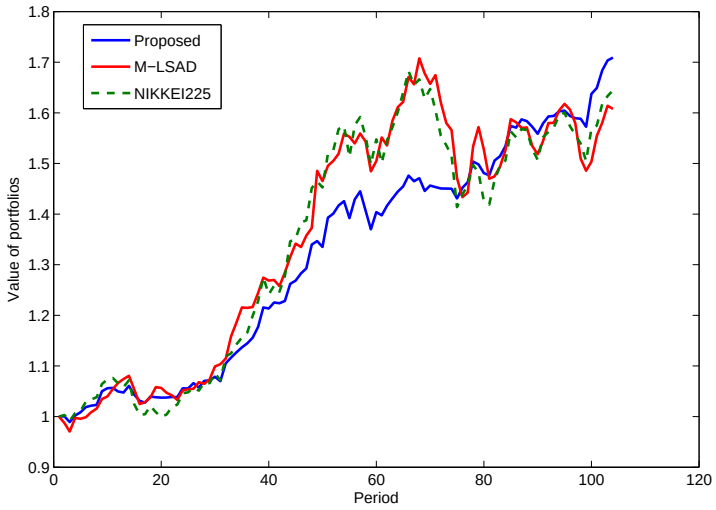

Figure 1 Fluctuation of the portfolios : Data 1

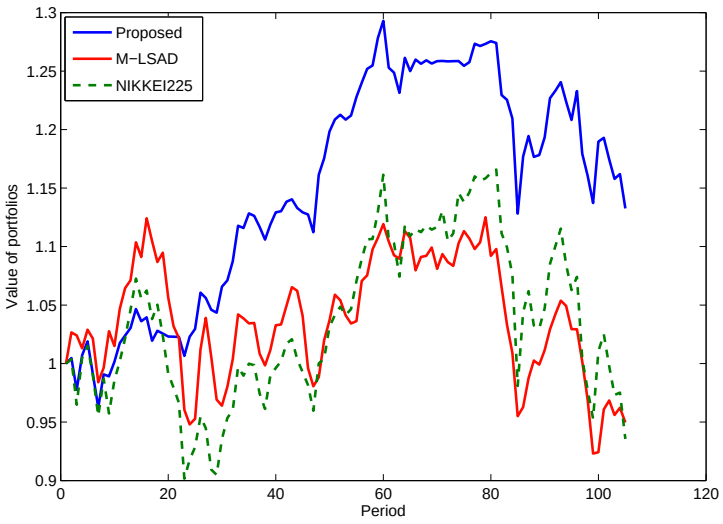

Figure 2 Fluctuation of the portfolios : Data 2

Moreover, Figures 3 and 4 show the expectation and the standard deviation of the obtained return rate with various required return rate γ for Data 1 and Data 2, respectively. From these figures, we see that

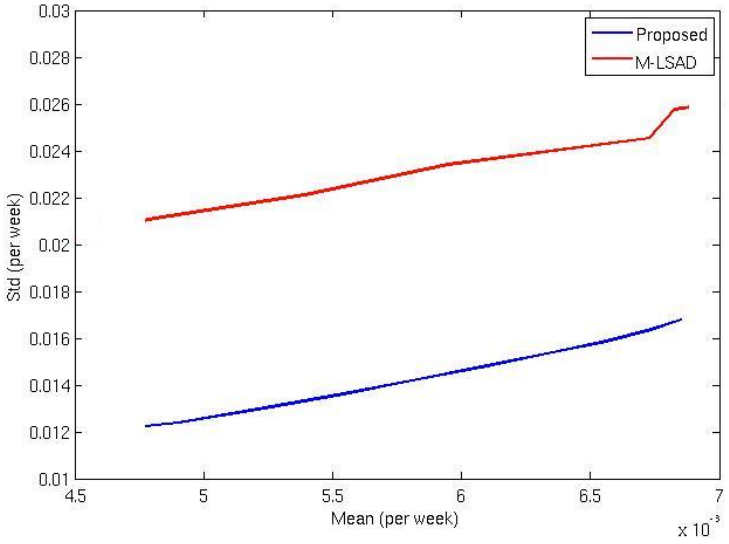

Figure 3 Relation between Mean and Std : Data 1

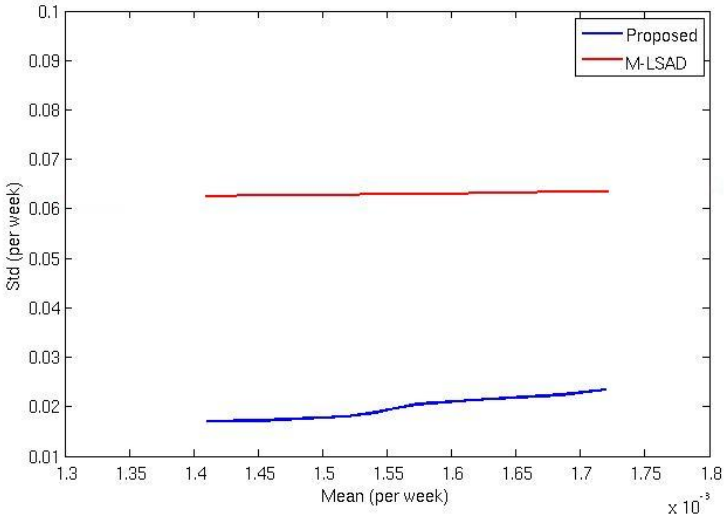

Figure 4 Relation between Mean and Std : Data 2

the risk by the proposed model is less than that by the M-LSAD model.

7 Conclusion

In this paper, we considered the portfolio function which outputs the optimal portfolio of the present economical situation, and formulated the optimization models to obtain the optimal portfolio function. We also derived the dual problem of the proposed models, and we showed that the optimal portfolio function is represented by the linear combinations of the kernel functions. From the numerical experiments, we see that the matrix splitting method can obtain the optimal portfolio function even if the number

of the assets is large. Moreover, we showed the validity of the proposed model as compared it with the M-LSAD model.

Our next step is to consider a transaction cost. If we change the portfolio frequently according to the situation, we cannot ignore the transaction cost. Hence we must consider the optimization models for the portfolio function with the transaction constraint.

Acknowledgements

First of all, I would like to express sincere appreciation to Associate Professor Nobuo Yamashita. His invaluable suggestion for my research encourage me. I also would like to tender my acknowledgements to Processor Masao Fukushima. I received the countless helpful comments from him. I am deeply indebted to Assistant Professor Shunsuke Hayashi. He gave me a lot of basic and applied knowledge. Moreover, I am very grateful to all members of Fukushima Laboratory, in particular, my peer, Yoshihiko Ito, Shota Yamanaka, Ke Ruan and Shunsuke Ozoe, for warm friendship. Finally, I would like to express may thanks to my father Akio Yoshida and my mother Masako Yoshida for their heartwarming kindness.

Reference

[1] D. P. Bertsekas: *Nonlinear Programming*, Athena Scientific, Belmont, 1995.

[2] C. M. Bishop: *Pattern Recognition and Machine Learing*, Springer, New York, 2006.

[3] R. W. Cottle, J.-S. Pang and R. E. Stone: *The Linear Complementarity Problem*, ACADEMIC PRESS, New York, 1992.

[4] H. Konno, H. Waki and A. Yuuki, *Portfolio Optimization under Lower Partial Risk Measures*, Asia-Pacific Financial Markets, **9** (2002), pp. 127–140.

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

[6] H. Markowitz, *Portfolio Analysis with Factors and Scenarios*, Journal of Finance, **36** (1959), pp. 871–877.

[7] H. M. Markowitz: *Portfolio Selection: Efficinet Diversification of Investments*, John Wiley & Sons, New York, 1959.

[8] Y. Nesterov and A. Nemirovsky: *Interior-Point Polynomial Algorithms in Convex Programming*, SIAM, Philadelphia, 1987.

[9] A. Perold, *Large Scale Portfolio Optimization*, Management Science, **30** (1984), pp. 1143–1160.

[10] T. R. Rockafellar and S. Uryasev, *Conditional Value-at-Risk for General Loss Distributions*, Jaurnal of Banking and Finance, **26** (2002), pp. 1443–1471.

[11] B. Schölkopf and A. J. Smola: *Learning with Kernels: Support Vector Machines, Regularization, Optimization and Beyond*, MIT Press, Cambridge, MA, 2002.

[12] W. F. Sharpe, *A Simplified Model for Portfolio Anaysis*, Management Science, **9** (1963), pp. 277–293.

[13] P. Tseng and S. Yun, *A Coordinate Gradient Descent Method for Linearly Constrainted Smooth Optimization and Suppot Vector Machine Training*, Computational Optimization and Applications, (2010), pp. 1–28.

[14] Yahoo! Finance, http://quote.yahoo.com.

A Data for Numerical Experiments

Table 6 shows the stocks which are used in numerical experiments.

Table 6 List of stocks used in numerical experiments

Stockname	code	Stockname	code
Ajinomoto	2802	Danippon Sumitomo Pharma	4506
Asahi Breweries	2502	Eisai	4523
Japan Tobacco	2914	Kyowa Hakko Holdings	4151
Kikkoman	2801	Shionogi &	4507
Kirin Brewery	2503	Takeda Pharmaceutical Company	4502
Nichirei	2871	Astellas Pharma	4503
Nippon Meat Packers	2282	Chugai Pharmaceutical	4519
Nisshin Seihun Group	2002	Showa Shell Sekiyu K.K.	5002
Sapporo Holdings	2501	The Yokohama Rubber	5101
Takara Holdings	2531	Bridgestone	5108
Unitika	3103	Nippon Sheet Glass	5202
Nisshinbo Industries	3105	Nitto Boseki	3110
Teijin	3401	Sumitomo Osaka Cement	5232
Toray Industries	3402	Taiheiyo Cement	5233
Toyobo	3101	Tokai Carbon	5301
Ojipaper	3861	NGK Insulators	5333
Hokuetsu Paper Mills	3865	Asahi Glass	5201
Mitsubishi Paper Mills	3864	Toto	5332
Ube Industries	4208	Sumitomo Metal Industries	5405
Asahi Kasei	3407	Pacific Metals	5541
Denki Kagaku Kogyo K.K.	4061	Nippon Steel	5401
Kuraray	3405	Kobe Steel	5406
Fujifilm Holdings	4901	Nisshin Steel	5407
Kao	4452	Toho Zinc	5707
Shin-Etsu Chemical	4063	The Furukawa Electric	5801
Mitsui Chemicals	4183	Mitsubishi Materials	5711
Nippon Kayaku	4272	Mitsui Mining & Smelting	5706
Nippon Soda	4041	Nippon Light Metal	5701
Nissan Chemical Industries	4021	Sumitomo Electric Industries	5802
Shiseido	4911	Sumitomo Metal Mining	5713
Showadenko K.K.	4004	Furukawa	5715
Sumitomo Chemical	4005	Fujikura	5803
Tosoh	4042	Dowa Holdings	5714

Stockname	code	Stockname	code
Sumitomo Heavy Industries	6302	Sharp	6753
Toyo Seikan Kaisya	5901	Sony	6758
Chiyoda	6366	Taiyo Yuden	6976
Daikin Industries	6367	TDK	6762
Ebara	6361	Tokyo Electron	8035
Hitachi Costruction Machinery	6305	Yokogawa Electric	6841
Mitsubishi Heavy Industries	7011	Kawasaki Heavy Industries	7012
NSK	6471	Mitsui Engineering & Shipbuilding	7003
NTN	6472	Ricoh	7752
Okuma Holdings	6103	Fuji Heavy Industries	7270
Kubota	6326	Hino Motors	7205
Komatsu	6301	Honda Motor	7267
JTEKT	6473	Isuzu Motors	7202
The Japan Steel Works	5631	Mazda Motor	7261
IHI	7013	Mitsubishi Motors	7211
Hitachi Zo-sen Corporation	7004	Nissan Motor	7201
Fanuc	6954	Suzuki Motor	7269
Fuji Electric Holdings	6504	Toyota Motor	7203
Fujitsu	6702	Citizen Holdings	7762
Kyocera	6971	Konica Minolta Holdings	4902
Hitachi	6501	Nikon	7731
Panasonic	6752	Olympus	7733
Panasonic Electric Works	6991	Terumo	4543
Meidensha	6508	Dai Nippon Printing	7912
Minebea	6479	Yamaha	7951
Mitsubishi Electric	6503	Toppan Printing	7911
Mitsumi Electric	6767	Nippon Suisan Kaisha	1332
NEC	6701	Daiwa House Industry	1925
Oki Electric Industry	6703	JGC	1963
Pioneer Corporation	6773	Kajima	1812
Denso	6902	Obayashi	1802
Casio Computer	6952	Sekisui House	1928
Canon	7751	Shimizu	1803
Alps Electric	6770	Taisei	1801
Advantest	6857	Itochu	8001
Toshiba	6502	Marubeni	8002
Sanyo Electric	6764	Mitsubishi	8058

Stockname	code	Stockname	code
Mitsui	8031	Central Japan Railway Company	9022
Sumitomo	8053	Tobu Railway	9001
Toyota Tsusho	8015	Nippon Express	9062
Aeon	8267	Yamato Holdings	9064
Fast Retailing	9983	Kawasaki Kisen Kaisha	9107
Uny	8270	Mitsui O.S.K. Lines	9104
Takashimaya	8233	Nippon Yusen K.K.	9101
The Chiba Bank	8331	All Nippon Airways	9202
The Bank of Yokohama	8332	Mitsubishi Logistics	9301
Mizuho Trust & Banking	8404	Softbank	9984
The Shizuoka Bank	8355	KDDI	9433
Nomura Holdings	8604	Nippon Telegraph and Telephone	9432
Daiwa Securities Group	8601	NTT Data	9613
Mizuho Securities	8606	NTT Docomo	9437
Credit Saison	8253	Chubu Electric Power	9502
Heiwa Real Estate	8803	The Kansai Electric Power	9503
Mitsubishi Estate	8802	The Tokyo Electric Power	9501
Mitsui Fudosan	8801	Osaka Gas	9532
Sumitomo Realty & Development	8830	Tokyo Gas	9531
Tokyu Land	8815	CSK Holdings	9737
East Japan Railway Company	9020	Konami	9766
Keio	9008	Secom	9735
Tokyu	9005	Tokyo Dome	9681
West Japan Railway Company	9021	Toho	9602
Keisei Electric Railway	9009	Yahoo Japan	4689
Odakyu Electric Railway	9007		