DISTRIBUTIONALLY ROBUST MULTI-PERIOD PORTFOLIO SELECTION SUBJECT TO BANKRUPTCY CONSTRAINTS

LIN JIANG AND CHANGZHI WU*

School of Management
Guangzhou University, Guangzhou, 400065, China

Song Wang

Department of Mathematics and Statistics
Curtin University, GPO Box U1987, Perth, WA 6845, Australia

(Communicated by Changjun Yu)

Abstract. An optimization problem with moments information which suffers from distributional uncertainty can be handled through distributionally robust optimization. In this paper, we will consider distributionally robust multi-period portfolio selection since only moment information of portfolios can be gathered in practice. We will consider two different scenarios. One is that moments information can be obtained exactly and the other one is that the moments information is also uncertain. For the two scenarios, we will show how to transform the corresponding distributionally robust optimization problem into a second order cone problem (SOCP) which can be easily solved by existing methods. Some numerical experiments are presented to demonstrate the effectiveness of our proposed method.

1. Introduction. Mean-variance portfolio selection was proposed by Markowitz in his seminal work [16] which was for single-period investment model. In this model, the return of investment measured by the mean of wealth is maximized while the risk measured by variance of portfolios to be selected is minimized. This model is then extended to the multi-period case [11]. In a multi-period mean-variance model, if the constraints are simple, this problem can be solved analytically. For example, analytical optimal portfolio policy and analytical expression of the mean-variance efficient frontier were derived in [11] through introducing an auxiliary parametric formulations which was to overcome nonseparable of the original problem in the sense of dynamic programming. Based on this technique, several multi-period portfolio selection problems are discussed, including the case with no shorting constraints [5], stochastic interest rate [20]. Instead of using auxiliary parametric formulations to tackle the issue of non-separability, a mean-field framework is introduced to tackle directly the issue of non-separability and derive the optimal policies analytically in [7]. Unfortunately, the analytical solution for a multi-period mean-variance model can be obtained only for those structurally simple cases.

2020 Mathematics Subject Classification. Primary: 90C26, 90C59; Secondary: 30E1.
Key words and phrases. Distributionally robust, multi-period portfolio selection, uncertainty, chance constraints.

*Corresponding author: Changzhi Wu (changzhiwu@gzhu.edu.cn).
Variance as a risk measure has been widely criticized by practitioners as it equally weights desirable positive returns and undesirable negative ones [15]. To circumvent this drawback, the semi-variance risk measure which only measures the variability of returns below the mean is introduced to replace variance [12]. Another typical kind of risk measures are Value-at-Risk (VaR) and Conditional-Value-at-Risk (CVaR). VaR is the quantile of the loss at a specified confidence level while CVaR is the conditional expected value of loss exceeding the VaR [6]. However, as a measure of risk, VaR lacks sub-additivity and convexity which leads the corresponding portfolio selection problem to being non-convexity [13]. To overcome the shortcomings of VaR, CVaR is proposed by Rockfellar and Uryasev in [17] which is proved to be coherent and convex. VaR and CVaR have been widely used in portfolio selection for both single-period and multi-period cases [15, 6]. In [15], single-period portfolio selection with mean-VaR model is studied which is a non-convex NP-hard optimization problem and an evolutionary based algorithm is proposed to solve the problem. In [2], a single period portfolio selection with covariance uncertainty is studied where Wasserstein distance is introduced to measure the covariance uncertainty. Under the weight constraints, the original problem is equivalent to a tractable optimization problem. In [21], a non-parametric approach is introduced to estimate the density of the loss function which leads to a convex formulation of the original portfolio selection problem. In [6], time inconsistency in multi-period mean-CVaR model is studied and time-consistent and self-coordination strategies are proposed. The proposed time-consistent strategy is a piecewise linear function of the wealth level with parameters which can be obtained through solving a series of mixed-integer programming problems off line. The self-coordination strategy is formulated as a convex program with a quadratic constraint.

In practical investment, the investment return is highly relying on the market. The underlying return distribution parameters, such as its expectation and covariance, cannot be obtained exactly in advance. In the recent years, there is a large amount of work to address this issue which is called as distributionally robust optimization [4, 8, 18]. Distributionally robust optimization is to handle the distributional uncertainty in stochastic optimization problems where the worst case of objective function and/or constraints are optimized under the given moment information. This model has been widely applied in portfolio selection to estimate the worst case of the investment risk. For example, in [9], for the given bounds on the mean and covariance matrix of returns, the worst-case VaR model of portfolio selection problem is studied. Through duality analysis, the proposed model can be cast as a semi-definite programming (SDP) which can be easily solved by existing convex optimization tools. In [23], the worst-case CVaR in robust portfolio selection is studied. For some special cases, it has been proved that the original robust optimization problem can be transformed into an equivalent SOCP. In [8], a distributionally robust optimization problem with uncertain moment information is studied. Through a series of duality analysis, the original problem can still be cast as a SDP although the mean and covariance are both suffered from uncertainty. Numerical experiments show that the daily return obtained under this distributionally robust framework is more reliable while not sacrificing daily utility. In [3], tight bounds on the expected values of several risk measures are studied. It has been shown that a single-period portfolio selection problem without additional constraints can always be solved analytically if the disutility function is in the form of
lower partial moments (LPM), VaR or CVaR. For the multi-period portfolio selection problem, it has been shown that the problem can still be solved analytically in [14]. In [12], a robust multi-period portfolio selection under downside risk LPM with asymmetrically distributed uncertainty set is studied. A computationally tractable approximation approach is proposed to solve the original problem.

For existing multi-period portfolio selection problems, most of them have ignored constraints and thus, the original problem can be solved analytically through applying dynamical programming. However, investment portfolios are always required to satisfy various constraints which are determined by the investment strategy. For example, chance constraints are considered in [19] where the chance constraints are handled by the one-sided Chebyshev’s inequality. Clearly, this approximation is not tight and thus the optimality of the obtained solution cannot be guaranteed. To overcome this problem, we will study multi-period portfolio selection under distributionally robust framework. In our discussions, we will study both the cases where the mean and covariance of returns are either known exactly or suffered from uncertainty. We will show that this problem can also be transformed into an equivalent SOCP which can be easily solved. Then, some numerical experiments are presented to illustrate and compare our proposed methods with those existing.

2. Multi-period portfolio selection with bankruptcy constraints. We consider a financial market with \( n + 1 \) assets available to be invested which consist of one cash riskless asset and \( n \) risky assets in a time horizon with \( T \) time periods. Let the cash riskless asset be labeled as 0 and the \( n \) risky assets be labeled as 1, \( \cdots, n \). Here the time period can be any time unit in accordance with real applications. Let \( s_t \) be the deterministic return of the riskless asset at period \( t \) and \( e_i^t \) be the random return of the risky asset \( i, i = 1, \cdots, n \). Denote the vector \( \mathbf{e}_t = [e_1^t, \cdots, e_n^t]^T \) to be the collection of all risky returns and the excess return vector of risky assets \( \mathbf{p}_t = [p_1^t, \cdots, p_n^t] = [e_1^t - s_t, \cdots, e_n^t - s_t]^T \). In the following discussions, we assume that the vector \( \mathbf{e}_t, t = 0, 1, \cdots, T \), are statistically independent and the only information known is its first two unconditional moments, i.e., its mean \( \mathbb{E}(\mathbf{e}_t) = [\mathbb{E}(e_1^t), \cdots, \mathbb{E}(e_n^t)]^T \) and its \( n \times n \) positive definite covariance \( \text{Cov}(\mathbf{e}_t) = \mathbb{E}(\mathbf{e}_t\mathbf{e}_t^T) - \mathbb{E}(\mathbf{e}_t)\mathbb{E}(\mathbf{e}_t^T) \).

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

\[
x_{t+1} = \sum_{i=1}^{n} e_i^t u_i^t + \left( x_t - \sum_{i=1}^{n} u_i^t \right) s_t \\
= s_t x_t + \mathbf{p}_t^T \mathbf{u}_t, \quad t = 0, 1, \cdots, T - 1. \tag{1}
\]
If we use CVaR to measure the risk and seek to maximize the terminal wealth, a multi-period portfolio selection problem with bankruptcy constraints can be formulated as the following:

\[
\text{MWC: } \max \mathbb{E}(x_T) \\
\text{s.t. } x_{t+1} = s_t x_t + p_t^T u_t \\
\mathbb{P}_t(x_t \geq \underline{x}) \geq 1 - \varepsilon, \ t = 1, \cdots, T,
\]

(2)

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

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

(3)

In practice, to get the exact distribution of \( \mathbb{P}_t \) is impossible. To overcome this difficulty, most of existing results are replacing the probability constraints (2) by standard constraints through using Tchebycheff inequality [7]. Although this approximation leads to a easily solved problem, the solution obtained is usually too conservative. Different from current methods, we will formulate this problem as a robust optimization problem. Let \( \tilde{\mu}_t \) and \( \tilde{\Sigma}_t \) be the estimates of the mean and covariance of the random vector \( \mathbb{P}_t \) based on the historical data. If these estimates are accurate, we define

\[
\mathcal{P}_t^1 = \{ \mathbb{P}_t \in \mathcal{M} : \mathbb{E}_{\mathbb{P}_t} (\mathbb{P}_t) = \tilde{\mu}_t, \mathbb{E}_{\mathbb{P}_t} \left[ (\mathbb{P}_t - \mu_t)((\mathbb{P}_t - \mu_t)^T) \right] = \tilde{\Sigma}_t \}, \ t = 1, \cdots, T.
\]

(4)

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

\[
\mathcal{P}_t^2 = \left\{ \mathbb{P}_t \in \mathcal{M} : \begin{array}{c}
\mathbb{E}_{\mathbb{P}_t} (\mathbb{P}_t - \mu_t)^T \tilde{\Sigma}_t^{-1} (\mathbb{E}_{\mathbb{P}_t} (\mathbb{P}_t) - \mu_t) \leq \gamma_1, \\
\mathbb{E}_{\mathbb{P}_t} \left[ (\mathbb{P}_t - \mu_t)(\mathbb{P}_t - \mu_t)^T \right] \leq \gamma_2 \tilde{\Sigma}_t
\end{array} \right\}, \ t = 1, \cdots, T.
\]

(5)

The first constraint in (5) describes how the estimate \( \tilde{\mu}_t \) is close to \( \mathbb{E}_{\mathbb{P}_t} (\mathbb{P}_t) \) while the second constraint in (5) enforces the covariance estimate to be bound in a semidefinite cone defined by a matrix inequality. Instead of replacing the probability constraints by the inequalities obtained through using the Tchebycheff inequality, we consider the following distributionally robust portfolios selection model:

\[
\text{DRMWC: } \max \mathbb{E}(x_T) \\
\text{s.t. } x_{t+1} = s_t x_t + p_t^T u_t \\
\min_{\mathbb{P}_t \in \mathcal{P}_t} \mathbb{P}_t(x_t \geq \underline{x}) \geq 1 - \varepsilon, \ t = 1, \cdots, T
\]

(6)

\[
\min_{\mathbb{P}_t \in \mathcal{P}_t} \mathbb{P}_t \left( x_t - \sum_{i=1}^n u_t^i \geq 0 \right) \geq 1 - \varepsilon, \ t = 1, \cdots, T - 1.
\]

(7)

where \( \mathcal{P}_t \) can be either \( \mathcal{P}_t^1 \) or \( \mathcal{P}_t^2 \).
3. Deterministic tractable reformulation and computation. In this part, we will reformulate Problem DRMWC as an equivalent deterministic problem without chance constraints, which is computationally tractable. In light of (1) and the independence of \( p_t, t = 1, \ldots, T \), we have

\[
\mathbb{E}(x_t) = T_0^t x_0 + \sum_{i=1}^{t-1} T_i^t \mu_{i-1}^T u_{i-1} + \mu_{t-1}^T u_{t-1},
\]

(8)

\[
\text{Var}(x_t) = \sum_{i=1}^{t-1} (T_i^t)^2 \Sigma_{i-1} u_{i-1} + u_{t-1}^T \Sigma_{t-1} u_{t-1},
\]

(9)

where \( T_0^t = \prod_{j=0}^{t-1} s_j \), \( T_i^t = \prod_{j=i}^{t-1} s_j \), \( \mathbb{E}_{\mathcal{P}_t}(p_t) = \mu_t \) and \( \mathbb{E}_{\mathcal{P}_t} [(p_t - \mu_t)(p_t - \mu_t)^T] = \Sigma_t \) are the mean and covariance matrix of the random vector of \( p_t \), respectively, for ease of notation. Now we only need to transform the constraints (2) into equivalent deterministic formulations.

**Lemma 3.1.** If the estimates \( \hat{\mu}_t \) and \( \hat{\Sigma}_t \) are exactly known, i.e., \( \mu_t = \hat{\mu}_t \), \( \Sigma_t = \hat{\Sigma}_t \), and \( \mathcal{P}_t = \mathcal{P}_t^1 \), then for \( t = 1, \ldots, T \), the inequalities (6) and (7) are equivalent to the following inequalities:

\[
\sqrt{\frac{\varepsilon}{1-\varepsilon}} \left( \sum_{i=1}^{t-1} (T_i^t)^2 \Sigma_{i-1} u_{i-1} + u_{t-1}^T \Sigma_{t-1} u_{t-1} \right)^{\frac{1}{2}} + \left( \varepsilon - \left( T_0^t x_0 + \sum_{i=1}^{t-1} T_i^t \mu_{i-1}^T u_{i-1} + \mu_{t-1}^T u_{t-1} \right) \right) \leq 0.
\]

(10)

\[
\sqrt{\frac{\varepsilon}{1-\varepsilon}} \left( \sum_{i=1}^{t-1} (T_i^t)^2 \Sigma_{i-1} u_{i-1} + u_{t-1}^T \Sigma_{t-1} u_{t-1} \right)^{\frac{1}{2}} + \left( \sum_{i=1}^{n} u_i^T - \left( T_0^t x_0 + \sum_{i=1}^{t-1} T_i^t \mu_{i-1}^T u_{i-1} + \mu_{t-1}^T u_{t-1} \right) \right) \leq 0.
\]

(11)

**Proof.** We can rewrite the inequalities (6) as

\[
\min_{\mathcal{P}_t} \mathbb{P}_t (\varepsilon - x_t) \leq 0.
\]

In light of Theorem 2.2 in [24], we know that

\[
\min_{\mathcal{P}_t} \mathbb{P}_t (\varepsilon - x_t) \geq 1 - \varepsilon \iff \sup_{\mathcal{P}_t} \mathbb{P}_t - \text{CVaR}_\varepsilon (\varepsilon - x_t) \leq 0.
\]

By virtue of Lemma 2.2 in [3], we have

\[
\sup_{\mathcal{P}_t} \mathbb{P}_t - \text{CVaR}_\varepsilon (\varepsilon - x_t)
\]

\[
= \min_{\alpha} \sup_{\zeta \sim (\varepsilon - \mathbb{E}(x_t), \text{Var}(x_t))} \alpha + \frac{1}{1-\varepsilon} \mathbb{E} \left[ \left( -\alpha - \zeta \right)_+ \right]
\]

\[
= \min_{\alpha} \alpha + \frac{1}{1-\varepsilon} \sup_{\zeta \sim (\varepsilon - \mathbb{E}(x_t), \text{Var}(x_t))} \mathbb{E} \left[ \left( -\alpha - \zeta \right)_+ \right]
\]

\[
= \min_{\alpha} \left\{ \alpha + \frac{1}{2(1-\varepsilon)} \left[ \sqrt{\text{Var}(x_t) + (\varepsilon - \mathbb{E}(x_t) + \alpha)^2} - (\alpha + \varepsilon - \mathbb{E}(x_t)) \right] \right\}
\]

where \( \zeta \sim (\varepsilon - \mathbb{E}(x_t), \text{Var}(x_t)) \) means that \( \zeta \) follows a distribution with the mean being \( \varepsilon - \mathbb{E}(x_t) \) and the covariance being \( \text{Var}(x_t) \).
Define
\[ h_\varepsilon(\alpha) = \alpha + \frac{1}{2(1-\varepsilon)} \left[ \sqrt{\text{Var}(x_t) + (\bar{x} - \mathbb{E}(x_t) + \alpha)^2} - (\alpha + \bar{x} - \mathbb{E}(x_t)) \right]. \]

Clearly, \( h_\varepsilon(\alpha) \) is convex in terms of \( \alpha \). Let \( \frac{\partial h_\varepsilon(\alpha)}{\partial \alpha} = 0 \), we obtain
\[ \alpha^* = \frac{2\varepsilon - 1}{2\sqrt{\varepsilon(1-\varepsilon)}} \sqrt{\text{Var}(x_t) + \bar{x} - \mathbb{E}(x_t)}. \]

Substituting \( \alpha^* \) into \( h_\varepsilon(\alpha) \), we obtain
\[ \sup_{\mathbb{P}_t \in \mathbb{P}_t} \mathbb{P}_t - \text{CVaR}_\varepsilon(\bar{x} - x_t) \]
\[ = \sqrt{\frac{\varepsilon}{1-\varepsilon}} \left( \sum_{i=1}^{t-1} (T^i_i)^2 u^T_{i-1} \bar{\Sigma}_{i-1} u_{i-1} + u^T_{t-1} \bar{\Sigma}_{t-1} u_{t-1} \right)^{\frac{1}{2}} \]
\[ - \left( \bar{x} - \left( T^0_0 x_0 + \sum_{i=1}^{t-1} T^i_i \mu_{i-1} u_{i-1} + \mu_{t-1} u_{t-1} \right) \right) \]

Similarly, we can prove the inequalities (11). The proof is completed. \( \square \)

Combining Lemma 3.1 and the equality (8) yields the following theorem:

**Theorem 3.2.** Problem DRMWC with \( \mathbb{P}_t = \mathbb{P}_t^1 \) is equivalent to the following optimization problem:

\[
\max T^T_0 x_0 + \sum_{i=1}^{T-1} T^T_i \mu_{i-1}^T u_{i-1} + \mu_{T-1}^T u_{T-1}
\]

s.t.
\[
\sqrt{\frac{\varepsilon}{1-\varepsilon}} \left( \sum_{i=1}^{t-1} (T^i_i)^2 u^T_{i-1} \Sigma_{i-1} u_{i-1} + u^T_{t-1} \Sigma_{t-1} u_{t-1} \right)^{\frac{1}{2}}
\]
\[+ \left( \bar{x} - \left( T^0_0 x_0 + \sum_{i=1}^{t-1} T^i_i \mu_{i-1}^T u_{i-1} + \mu_{t-1}^T u_{t-1} \right) \right) \leq 0, \ \text{for} \ \ t = 1, \cdots, T, (13)\]

\[
\sqrt{\frac{\varepsilon}{1-\varepsilon}} \left( \sum_{i=1}^{t-1} (T^i_i)^2 u^T_{i-1} \Sigma_{i-1} u_{i-1} + u^T_{t-1} \Sigma_{t-1} u_{t-1} \right)^{\frac{1}{2}}
\]
\[+ \left( \sum_{i=1}^{n} u^i_i - \left( T^0_0 x_0 + \sum_{i=1}^{t-1} T^i_i \mu_{i-1}^T u_{i-1} + \mu_{t-1}^T u_{t-1} \right) \right) \leq 0, \ \text{for} \ \ t = 1, \cdots, T - 1. \]

(14)

Based on Theorem 3.2, Problem DRMWC has been transformed into an equivalent SOCP problem which can be easily solved. Now we study the case \( \mathbb{P}_t = \mathbb{P}_t^2 \). In the following discussions, we further assume that all the wealth is invested in the risky market without cash keeping. Thus, the dynamics of the wealth (1) becomes
\[ x_{t+1} = \sum_{i=1}^{n} e^i_{t} u^i_{t} = e^T_{t} u_{t}. \]
Under this environment, Problem DRMWC should be proposed as follows:

\[
\text{DRMWC: } \max_{u} \inf_{P \in P_t^i} \mathbb{E}(x_T) \tag{16}
\]

\[
\text{s.t. } x_{t+1} = e_{t+1}'u_t \\
\inf_{P \in P_t^i} \mathbb{P}_t(x_t \geq z) \geq 1 - \varepsilon, \quad t = 1, \cdots, T, \tag{17}
\]

\[
\inf_{P \in P_t^i} \mathbb{P}_t(1^Tu_t - e_{t}^Tu_{t-1} \geq 0) \geq 1 - \varepsilon, \quad t = 1, \cdots, T - 1. \tag{18}
\]

Clearly, solving Problem DRMWC with \(P_t = P_t^i\) is more difficult than \(P_t = P_t^1\) if \(\mu_t\) and \(\Sigma_t\) cannot be accessed perfectly. To circumvent this difficulty, we need to decompose Problem DRMWC with \(P_t = P_t^2\) as a two layer optimization problem in which the inner layer is the problem with \(P_t = P_t^1\) while the outer layer is to handle estimation inaccuracy of \(\mu_t\) and \(\Sigma_t\). To further our discussion, we need the following lemma which is a variation of Lemma 3.3 in [22]:

**Lemma 3.3.** Let \(s\) be a random vector in \(\mathbb{R}^d\) and let \(\xi\) be a random variable in \(\mathbb{R}\). For a given \(y \in \mathbb{R}^d\), let ambiguity sets \(D_s\) and \(D_\xi\) be as follows:

\[
D_s = \{ P \in \mathcal{P}(\mathbb{R}^d) : \mathbb{E}_P(s)^T\Sigma^{-1}_s\mathbb{E}_P(s) \leq \gamma_1, \quad \mathbb{E}_P(ss^T) \leq \gamma_2\Sigma \}
\]

\[
D_\xi = \{ P \in \mathcal{P}(\mathbb{R}) : \|\mathbb{E}_P(\xi)\| \leq \sqrt{\gamma_1 y^T\Sigma y}, \quad \mathbb{E}_P(\xi^2) \leq \gamma_2 y^T\Sigma y \}
\]

Then, we have

\[
\inf_{P \in D_s} \mathbb{E}_P(y^Ts) = \inf_{P \in D_\xi} \mathbb{E}_P(\xi).
\]

Furthermore,

\[
\inf_{P \in D_s} \mathbb{E}_P(y^s) = \begin{cases} -\sqrt{\gamma_1 y^T\Sigma y}, & \text{if } \gamma_1 \leq \gamma_2, \\ -\sqrt{\gamma_2 y^T\Sigma y}, & \text{otherwise.} \end{cases} \tag{19}
\]

**Proof.** For any \(P \in D_s\), define \(\xi = y^Ts\). It yields that

\[
\mathbb{E}_P(\xi) = \mathbb{E}_P(y^Ts) = y^T\mathbb{E}_P(s) \leq \|\Sigma^{-1}_s\|\|y\|^2\mathbb{E}_P(s)\|.
\]

Since \(s \in D_s\), \(\|\Sigma^{-1}_s\|\mathbb{E}_P(s)\| \leq \sqrt{\gamma_1}\). Replacing this inequality into (20) yields that

\[
\mathbb{E}_P(\xi) \leq \sqrt{\gamma_1}\|\Sigma^{-1}_s\|\|y\| = \sqrt{\gamma_1} y^T\Sigma y.
\]

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

\[
\mathbb{E}_P(\xi) \geq -\sqrt{\gamma_1}\|\Sigma^{-1}_s\|\|y\| = -\sqrt{\gamma_1} y^T\Sigma y.
\]

Thus, for any \(s \in D_s\) we have \(\xi = y^Ts \in D_\xi\). It implies that

\[
\inf_{P \in D_s} \mathbb{E}_P(y^Ts) \leq \inf_{P \in D_\xi} \mathbb{E}_P(\xi).
\]

On the contrary, for any \(\xi \in D_\xi\), define \(s = \xi\Sigma y/y^T\Sigma y\), then

\[
\mathbb{E}_P(s)^T\Sigma^{-1}_s\mathbb{E}_P(s) = (\mathbb{E}_P(\xi))^2 y^T\Sigma\Sigma^{-1}_s\Sigma y/(y^T\Sigma y)^2 = \frac{(\mathbb{E}_P(\xi))^2}{y^T\Sigma y} \leq \gamma_1.
\]

Meanwhile,

\[
\mathbb{E}_P(ss^T) = \mathbb{E}_P\left\{ \xi^2 \Sigma y/y^T\Sigma y \right\} = \mathbb{E}_P(\xi^2) \left\{ \frac{\Sigma y}{y^T\Sigma y} \right\} \leq \gamma_2 y^T\Sigma y/(y^T\Sigma y)^2.
\]

Furthermore, we claim that \(\Sigma y/(y^T\Sigma y) \leq y^T\Sigma y\), we have \(\mathbb{E}_P(ss^T) \leq \gamma_2\Sigma\). Indeed, for any \(z \in \mathbb{R}^d\), we have

\[
\mathbb{E}_P(ss^T) = \mathbb{E}_P\left\{ \xi^2 \Sigma y/y^T\Sigma y \right\} = \mathbb{E}_P(\xi^2) \leq \gamma_2 y^T\Sigma y/(y^T\Sigma y)^2.
\]
Thus, the following inequality holds:

\[ E_P(\xi^2) = y^T E_P(ss^T)y \leq \gamma_2 y^T \Sigma y. \]

Therefore, \( s \in D_s \) and

\[
\inf_{P \in \mathcal{D}_s} E_P(y^T s) \geq \inf_{P \in \mathcal{D}_\xi} E_P(\xi).
\]

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

\[ E_P(\xi) \geq -\sqrt{\gamma_2 y^T \Sigma y}. \]

The inequality can also be attained. Therefore,

\[
\inf_{P \in \mathcal{D}_s} E_P(y^T s) = \begin{cases} 
-\sqrt{\gamma_1} y^T \Sigma y, & \text{if } \gamma_1 \leq \gamma_2, \\
-\sqrt{\gamma_2} y^T \Sigma y, & \text{otherwise}.
\end{cases}
\]

We complete the proof.

To proceed it further, we cite Theorem 3.2 in [22] as the following lemma (Lemma 3.4) which will be used later.

**Lemma 3.4.** Suppose that

\[
\mathcal{D} = \{ P \in \mathcal{P}(\mathbb{R}^J) : E_P(s)^T \Sigma^{-1} E_P(s) \leq \gamma_1, E_P(ss^T) \leq \gamma_2 \Sigma \} \tag{22}
\]

Then, \( \inf_{P \in \mathcal{D}} \{ s^T y \leq M \} \geq 1 - \alpha \) is equivalent to

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

if \( \gamma_1 / \gamma_2 \leq \alpha \), and is equivalent to

\[
\mu^T y + \frac{\gamma_2}{\alpha} \sqrt{y^T \Sigma y} \leq T
\]

if \( \gamma_1 / \gamma_2 > \alpha \).

Now, we can show that the distributional robust multi-period portfolio selection problem with uncertain moments \( \mathcal{P}_t = \mathcal{P}_t^2 \) can be transformed into a SOCP. From this point, Problem DRMWC with \( \mathcal{P}_t = \mathcal{P}_t^2 \) has the same computational complexity as that of \( \mathcal{P}_t = \mathcal{P}_t^1 \). It means that the uncertainty of the moments does not increase the complexity of the problem.
Theorem 3.5. If $\gamma_1 \leq \gamma_2$, Problem DRMWC with $\mathcal{P}_t = \mathcal{P}^2_t$ is equivalent to the following SOCP:

$$\begin{align*}
\max_{\bar{u}_t} & \quad \bar{\mu}^T_T u_{T-1} - \sqrt{\gamma_1 \mu^T_{T-1} \Sigma_T \mu_{T-1}} \\
\text{s.t.} & \quad \left( \sqrt{\gamma_1} + \frac{1-\varepsilon}{\varepsilon} (\gamma_2 - \gamma_1) \right) \sqrt{\mu^T \Sigma \mu} \leq \bar{\mu}^T T u_{t-1} - \bar{x}, \quad t = 1, \cdots, T,
\end{align*}$$

(23)

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

$$\begin{align*}
\max_{\bar{u}_t, \gamma} & \quad \bar{\mu}^T_T u_{T-1} - \sqrt{\gamma_2 \mu^T_{T-1} \Sigma_T \mu_{T-1}} \\
\text{s.t.} & \quad \left( \sqrt{\gamma_1} + \frac{1-\varepsilon}{\varepsilon} (\gamma_2 - \gamma_1) \right) \sqrt{\mu^T \Sigma \mu} \leq \bar{\mu}^T T u_{t-1} - \bar{x}, \quad t = 1, \cdots, T,
\end{align*}$$

(24)

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

$$\begin{align*}
\max_{\bar{u}_t, \gamma} & \quad \bar{\mu}^T_T u_{T-1} - \sqrt{\gamma_2 \mu^T_{T-1} \Sigma_T \mu_{T-1}} \\
\text{s.t.} & \quad \sqrt{\gamma_1} \sqrt{\mu^T \Sigma \mu} \leq \bar{\mu}^T T u_{t-1} - \bar{x}, \quad t = 1, \cdots, T;
\end{align*}$$

(25)

Proof. Denote $s_t = e_t - \bar{\mu}_t$. Then, we can verify that

$$\mathcal{D}_s = \{ \mathbb{P} \in \mathcal{P}^{J} : \mathbb{E}_\mathbb{P}(s_t) \Sigma^{-1} \mathbb{E}_\mathbb{P}(s_t) \leq \gamma_1, \mathbb{E}_\mathbb{P}(s_t s_t^T) \leq \gamma_2 \Sigma \}$$

In light of Lemma 3.3, we have

$$\begin{align*}
\inf_{\mathbb{P} \in \mathcal{D}_s} \mathbb{E}_\mathbb{P}(s_t^T u_{T-1}) & = \begin{cases} 
-\sqrt{\gamma_1} \sqrt{\mu^T_{T-1} \Sigma_T \mu_{T-1}}, & \text{if } \gamma_1 \leq \gamma_2, \\
-\sqrt{\gamma_2} \sqrt{\mu^T_{T-1} \Sigma_T \mu_{T-1}}, & \text{otherwise}.
\end{cases} \\
\end{align*}$$

(32)

Since $x_T = s_T^T u_{T-1} + \bar{\mu}^T T u_{T-1}$,

$$\begin{align*}
\inf_{\mathbb{P} \in \mathcal{D}_s} \mathbb{E}_\mathbb{P}(x_T) & = \begin{cases} 
\bar{\mu}^T T u_{T-1} - \sqrt{\gamma_1} \sqrt{\mu^T_{T-1} \Sigma_T \mu_{T-1}}, & \text{if } \gamma_1 \leq \gamma_2, \\
\bar{\mu}^T T u_{T-1} - \sqrt{\gamma_2} \sqrt{\mu^T_{T-1} \Sigma_T \mu_{T-1}}, & \text{otherwise}.
\end{cases}
\end{align*}$$

(33)
In a similar way, we can prove that
\[
\inf_{P_t \in P_t^1} \mathbb{P}_t(x_t \geq x) \geq 1 - \varepsilon \iff
\begin{align*}
\left\{ \begin{array}{l}
\left( \sqrt{\gamma_1} + \sqrt{\frac{1-\varepsilon}{\varepsilon} (\gamma_2 - \gamma_1)} \right) \sqrt{u_t^T \Sigma_t u_t} \\
\leq \hat{\mu}_t^T u_{t-1} - \frac{x}{\alpha}, \text{ if } \gamma_1/\gamma_2 \leq \alpha,
\end{array} \right.
\end{align*}
\]
(34)
inf_{P_t \in P_t^1} \mathbb{P}_t (u_t - e_t^T u_{t-1} \geq 0) \iff
\begin{align*}
\left\{ \begin{array}{l}
\hat{\mu}_{t-1}^T u_{t-1} + \left( \sqrt{\gamma_1} + \sqrt{\frac{1-\varepsilon}{\varepsilon} (\gamma_2 - \gamma_1)} \right) \sqrt{u_{t-1}^T \Sigma_{t-1} u_{t-1}} \\
\leq \sum_{i=1}^n u_t^i, \text{ if } \gamma_1/\gamma_2 \leq \alpha,
\end{array} \right.
\end{align*}
(35)
Then, the results (23) - (31) are easily obtained through combining the above results. We completed the proof. \qed

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

Example 1. Let us first consider the case without moment uncertainties. For this case, we select three portfolios from Shanghai Stock exchange. The transaction data within 60 business days is used to compute \( \mu_t \) and \( \Sigma_t \). Let \( T = 10 \) and \( s_t = 1.04 \). We suppose that the mean and covariance are constant during this time period. Then, the corresponding \( \mu_t = [0.122, 0.206, 0.188]^T \) and
\[
\Sigma_t = \begin{bmatrix}
0.0146 & 0.0187 & 0.0145 \\
0.0187 & 0.0854 & 0.0104 \\
0.0145 & 0.0104 & 0.0289
\end{bmatrix}
\]
Let \( \bar{x} = 1.15 \) and \( \varepsilon = 0.05 \). Then, the expected return at the end of the period is \( \mathbb{E}(x_T) = 8.0758 \). If we adjust \( \bar{x} \) from \( \bar{x} = 1.15 \) to \( \bar{x} = 1.196 \), then the expected return \( \mathbb{E}(x_T) = 5.3168 \). The optimal \( \mathbb{E}(x_t) \) with \( \bar{x} = 1.15 \) and \( \bar{x} = 1.196 \) are depicted in Figure 1. From Figure 1, we can observe clearly that the increase of \( \bar{x} \) has a significant decrease of the expected return \( \mathbb{E}(x_T) \). If we set \( \bar{x} = 1.2 \), then no feasible solution is found.

Fig 2 and Fig 3 show the optimal \( u_2(t) \) and \( u_3(t) \) with \( \bar{x} = 1.15 \) and \( \bar{x} = 1.196 \). From Fig 2 and Fig 3, we can observe that the smaller \( \bar{x} \), the larger \( u_2(t) \) and the smaller \( u_3(t) \). The reason behind is that the second portfolio has the largest investment return, but it has the largest risk from the variance perspective. The expected investment return of the third portfolio is between the first one and the second one. The increase of \( \bar{x} \) means the aversion of the risk. Thus, \( u_3(t) \) will be increased with the increase of \( \bar{x} \) in order to reduce the risk from the investment of the second portfolios.

Example 2. Now we consider the case with uncertain moments. The expected investment return and the variance matrix are still the same as those in Example 1. The problem defined by (16)-(18) is different from the problem defined by (6)-(7) as \( x_t \) is expressed only in terms of \( u_t \) in the problem with uncertain moments. This expression is adopted as the problem with uncertain moments can be transferred into an equivalent second order cone programming under this formulation. We further constrain that \( \sum_{i=1}^n u_i(t) = 1 \), for all \( t = 1, \cdots, T \) which will show the percentage of each portfolio invested at different times. In this case, the investment return becomes \( \mu_t = [0.162, 1.246, 1.228] \). Now let \( \gamma_1 = 0.0001 \) and \( \gamma_2 = 1.5 \). At
Figure 1. The optimal $E(x_t)$ with $x = 1.15$ and $x = 1.196$

Figure 2. The optimal $u_2(t)$ with $x = 1.15$ and $x = 1.196$

At the beginning, we suppose that the portfolios are equally distributed, i.e., $u_{0,1}(0) = u_{0,2}(0) = u_{0,3}(0) = 1/3$. If we set $x = 1$, then no feasible solution is found. Let $x = 0.85$. The obtained $u_i(t)$, $i = 1, 2, 3$, are depicted in Fig 4. From the figure, we can observe that at the beginning stage, the portions of the three portfolios are similar. However, the second portfolios will increase significantly with the time evolution. The reason is that the second portfolio has the largest investment return, but it also has the largest variance. With the increase of the time, the expected worst investment returns will be large than $x$ under the given moment uncertainty. Since the expected investment return is maximized, the portion of the second portfolio is of course becoming larger and larger.
Now let us vary the parameters $\gamma_1$ and $\gamma_2$ to observe its impact on the optimal solution. From the definition of $\mathcal{P}_T^2$ in (5), we can see that the parameters $\gamma_1$ and $\gamma_2$ regulate the boundary of the uncertainties. With the increase of the $\gamma_1$ and $\gamma_2$, the uncertainty set is increased and thus, the optimal investment return under the worst case uncertainty will be decreased. Set $\gamma_1 = 0.0001$ and vary $\gamma_2$, the corresponding optimal investment returns under the worst distribution scenario are shown in Fig 5. We can clearly observe that with the increase of the parameter $\gamma_2$, the investment return is decreased.

Now we fix $\gamma_2 = 1$ and vary $\gamma_1$. From Fig 6, we can see that the with the increase of $\gamma_1$, the optimal investment return is also decreased. If we set $\gamma_1 = 0.01$ with $\gamma_2 = 1$, then no feasible solution is found.
5. **Conclusion.** In this paper, we have studied the dynamic portfolio selection problem with distributional uncertainty. If the moments are known exactly, then this problem can be transformed into an equivalent second order cone programming. If the moments cannot be known exactly but within a norm bounded set, we can still prove that it can be transformed into a second order cone programming. Two simple numerical examples are presented to illustrate our proposed method. The numerical results are consistent with our intuition in practice.

**REFERENCES**

[1] S. Alexander, T. Coleman and Y. Li, Minimizing CVaR and VaR for a portfolio of derivatives, *Journal of Banking and Finance*, 30 (2006), 583–605.
[2] J. Blanchet, L. Chen and X. Zhou, Distributionally robust mean-variance portfolio selection with Wasserstein distances, arXiv:1802.02885.
[3] L. Chen, S. He and S. Zhang, Tight bounds for some risk measures, with applications to robust portfolio selection, Oper. Res., 59 (2011), 847–865.
[4] Y. Chen, H. Sun and H. Xu Decomposition and discrete approximation methods for solving two-stage distributionally robust optimization problems, Comput. Optim. Appl., 78 (2021), 205–238.
[5] X. Cui, J. Gao, X. Li and D. Li, Optimal multi-period mean–variance policy under no-shorting constraint, European J. Oper. Res., 234 (2014), 459–468.
[6] X. Cui, J. Gao, Y. Shi and S. Zhu, Time-consistent and self-coordination strategies for multi-period mean-conditional value-at-risk portfolio selection, European J. Oper. Res., 276 (2019), 781–789.
[7] X. Cui, X. Li and D. Li, Unified framework of mean-field formulations for optimal multi-period mean-variance portfolio selection, IEEE Trans. Automat. Control, 59 (2014), 1833–1844.
[8] E. Delage and Y. Ye, Distributionally robust optimization under moment uncertainty with application to data-driven problems, Oper. Res., 58 (2010), 595–612.
[9] L. Ghaoui, M. Oks and F. Oustry, Worst-case value-at-risk and robust portfolio optimization: A conic programming approach, Oper. Res., 51 (2003), 543–556.
[10] B. Li, Y. Zhu, Y. Sun, G. Aw and K. Teo, Multi-period portfolio selection problem under uncertain environment with bankruptcy constraint, Appl. Math. Model., 56 (2018), 539–550.
[11] D. Li and W. Ng, Optimal dynamic portfolio selection: Multiperiod mean-variance formulation, Math. Finance, 10 (2000), 387–406.
[12] A. Ling, J. Sun and M. Wang, Robust multi-period portfolio selection based on downside risk with asymmetrically distributed uncertainty set, European J. Oper. Res., 285 (2020), 81–95.
[13] B. Liu, A new risk measure and its application in portfolio optimization: The SPP–CVaR approach, Economic Modelling, 51 (2015), 383–390.
[14] J. Liu, Z. Chen, A. Lisser and Z. Xu, Closed-form optimal portfolios of distributionally robust mean-CVaR problems with unknown mean and variance, Appl. Math. Optim., 79 (2019), 671–693.
[15] K. Lwin, R. Qu and B. MacCarthy, Mean-Var portfolio optimization: A nonparametric approach, European J. Oper. Res., 260 (2017), 751–766.
[16] H. Markowitz, Portfolio Selection, Cowles Foundation for Research in Economics at Yale University, Monograph 16 John Wiley & Sons, Inc., New York; Chapman & Hall, Ltd., London 1959.
[17] R. T. Rockafellar and S. Uryasev, Optimization of conditional value-at-risk, Journal of Risk, 2 (2000), 21–42.
[18] H. Sun and H. Xu, Convergence analysis for distributionally robust optimization and equilibrium problems, Math. Oper. Res., 41 (2016), 377–401.
[19] Y. Sun, G. Aw, R. Loxton and K. L. Teo, Chance-constrained optimization for pension fund portfolios in the presence of default risk, European J. Oper. Res., 256 (2017), 205–214.
[20] H. Yao, Z. Li and D. Li, Multi-period mean-variance portfolio selection with stochastic interest rate and uncontrollable liability, European J. Oper. Res., 252 (2016), 837–851.
[21] H. Yao, Z. Li and Y. Lai, Mean–CVaR portfolio selection: A nonparametric estimation framework, Comput. Oper. Res., 40 (2013), 1014–1022.
[22] Y. Zhang, R. Jiang and S. Shen, Ambiguous chance-constrained binary programs under mean-covariance information, SIAM J. Optim., 28 (2018), 2922–2944.
[23] S. Zhu and M. Fukushima, Worst-case conditional value-at-risk with application to robust portfolio management, Oper. Res., 57 (2009), 1155–1168.
[24] S. Zymler, D. Kuhn and B. Rustem Distributionally robust joint chance constraints with second-order moment information, Math. Program., 137 (2013), 167–198.

Received April 2021; revised October 2021; early access December 2021.

E-mail address: lydianjianglin@gmail.com
E-mail address: Song.Wang@curtin.edu.au
E-mail address: changzhiwu@gzhu.edu.cn