Chebyshev Inequality based Approach to Chance Constrained Portfolio Optimization

KIYOHARU TAGAWA
Kindai University
School of Science and Engineering
3-4-1 Kowakae, Higashi-Osaka, 577-8502
JAPAN
tagawa@info.kindai.ac.jp

Abstract: A new approach to solve Chance constrained Portfolio Optimization Problems (CPOPs) without using the Monte Carlo simulation is proposed. Specifically, according to Chebyshev inequality, the prediction interval of a stochastic function value included in CPOP is estimated from a set of samples. By using the prediction interval, CPOP is transformed into Lower-bound Portfolio Optimization Problem (LPOP). It is proved that the feasible solution of LPOP is also feasible for CPOP. Furthermore, in order to solve LPOP, Differential Evolution (DE) is used. Finally, through a numerical experiment, the usefulness of the proposed approach is demonstrated.

Key–Words: Portfolio Optimization, Risk Management, Chance Constraint, Chebyshev Inequality

1 Introduction

The theory of portfolio optimization was provided by Markowitz [1]. The Markowitz model evaluates the risk of investment by the variance of its return. Thereby, the model takes a balance between the total return and the risk aversion by choosing their weights appropriately. Even though the Markowitz model is a standard formulation of portfolio optimization, there are some questions to be answered including the choice of the weights. Therefore, the Markowitz model has been greatly developed nowadays [2, 3].

Kataoka [2] formulated the portfolio selection as a chance constrained optimization problem in which the quantile criterion was introduced with a specified probability $\alpha \in (0, 1)$. The Kataoka model is called Chance constrained Portfolio Optimization Problem (CPOP) in this paper. The risk of investment can be limited by a specified probability in CPOP. However, CPOP is usually difficult to solve. That is because the probability that the chance constraint is satisfied has to be evaluated empirically by using a time-consuming Monte Carlo simulation [4]. Therefore, instead of the theory of probability, the fuzzy theory is used widely to formulate modified CPOPs [5, 6, 7, 8, 9].

A new technique has recently been proposed to solve chance constrained optimization problems without using the time-consuming Monte Carlo simulation [10]. This paper applies the new technique to the primitive CPOP. Specifically, by using Chebyshev inequality, CPOP is transformed into Lower-bound Portfolio Optimization Problem (LPOP). It is proved that the feasible solution of LPOP is also feasible for CPOP. Furthermore, one of the powerful evolutionary algorithms, namely Differential Evolution (DE) [11], is used to find the optimal solution of LPOP.

The rest of this paper is organized as follows. Section 2 describes CPOP briefly. According to the new approach, CPOP is transformed into LPOP in Section 3. Section 4 presents DE for solving LPOP. Section 5 shows the result of numerical experiments conducted on an instance of CPOP. Finally, Section 6 gives conclusion and future work.

2 Portfolio Optimization

We describe CPOP [2]. First of all, we assume that a portfolio is composed by D assets. Let $x_i \in \Re$, $i = 1, \cdots, D$ be a ratio of the i-asset. Therefore, a vector of decision variables of CPOP is represented as $\boldsymbol{x} = (x_1, \cdots, x_D) \in \boldsymbol{X}$, $\boldsymbol{X} = [0, 1]^D \subseteq \Re^D$.

The unit investment in the i-asset provides the random return ξ_i over a considered fixed period. We assume that the vector of random returns $\boldsymbol{\xi} \in \Re^D$ is characterized by a multivariate normal distribution

$$\boldsymbol{\xi} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{V}) \tag{1}$$

with mean $\boldsymbol{\mu} \in \Re^D$ and covariance matrix:

$$\boldsymbol{V} = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1D} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2D} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{D1} & \sigma_{D2} & \cdots & \sigma_{DD} \end{bmatrix}. \tag{2}$$

We try to maximize the expected total return $g(\boldsymbol{x}, \boldsymbol{\xi}) = \boldsymbol{\xi}^T \boldsymbol{x}$ for a given probability $\alpha \in (0, 1)$, which denotes the risk of the investment. By using an objective variable $\gamma \in \Re$, CPOP is formulated as

$$\left[\begin{array}{ll} \max\limits_{\boldsymbol{x} \in \boldsymbol{X}} & \gamma \\ \text{sub. to} & \Pr(g(\boldsymbol{x}, \boldsymbol{\xi}) \geq \gamma) \geq 1 - \alpha, \\ & x_1 + x_2 + \cdots + x_D = 1 \end{array}\right. \tag{3}$$

where $\Pr(e)$ denotes the probability of an event e.

3 Proposed Approach

3.1 Chebyshev Inequality

Due to $\boldsymbol{\xi} \in \Re^D$, different function values $g(\boldsymbol{x}, \boldsymbol{\xi})$ are observed in (3) for repeated evaluations of the same solution $\boldsymbol{x} \in \boldsymbol{X}$. If the mean $\mu(\boldsymbol{x})$ and the variance $\sigma^2(\boldsymbol{x})$ of the random variable $g(\boldsymbol{x}, \boldsymbol{\xi})$ was known, Chebyshev inequality [12] could be stated as

$$\Pr(|g(\boldsymbol{x}, \boldsymbol{\xi}) - \mu(\boldsymbol{x})| \geq \lambda\, \sigma(\boldsymbol{x})) \leq \frac{1}{\lambda^2} \tag{4}$$

where $\lambda > 1$ is an arbitrary real number.

Chebyshev inequality has great utility because it can be applied to completely arbitrary distributions. However, the values of $\mu(\boldsymbol{x})$ and $\sigma^2(\boldsymbol{x})$ used in (4) are usually unknown. Therefore, Saw et al. [13] have extended Chebyshev inequality in (4) to cases where the mean and variance are not known and may not exist, but you want to use the sample mean and sample variance from N samples to bound the expected value of a new drawing from the same distribution.

First of all, we take N samples of $g(\boldsymbol{x}, \boldsymbol{\xi})$ as

$$\begin{aligned} g(\boldsymbol{x}, \boldsymbol{\xi}^n) &= (\boldsymbol{\xi}^n)^T \boldsymbol{x} \\ &= \xi_1^n\, x_1 + \xi_2^n\, x_2 + \cdots + \xi_D^n\, x_D \end{aligned} \tag{5}$$

where $\boldsymbol{\xi}^n \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{V}),\ n = 1, 2, \cdots, N$.

From the set of samples in (5), the sample mean $\overline{g}(\boldsymbol{x})$ and the sample variance $s^2(\boldsymbol{x})$ are obtained as

$$\overline{g}(\boldsymbol{x}) = \frac{1}{N} \sum_{n=1}^{N} g(\boldsymbol{x}, \boldsymbol{\xi}^n), \tag{6}$$

$$s^2(\boldsymbol{x}) = \frac{1}{N-1} \sum_{n=1}^{N} (g(\boldsymbol{x}, \boldsymbol{\xi}^n) - \overline{g}(\boldsymbol{x}))^2. \tag{7}$$

By using $\overline{g}(\boldsymbol{x})$ and $s^2(\boldsymbol{x})$, Chebyshev inequality from samples [13] is stated as

$$\begin{aligned} \Pr\left(|g(\boldsymbol{x}, \boldsymbol{\xi}) - \overline{g}(\boldsymbol{x})| \geq \lambda \sqrt{\frac{N+1}{N}}\, s(\boldsymbol{x})\right) \\ \leq \frac{1}{N+1} \left\lfloor \frac{(N+1)\,(N-1+\lambda^2)}{N\,\lambda^2} \right\rfloor \end{aligned} \tag{8}$$

where $\lfloor r \rfloor$ denotes the floor function of $r \in \Re$.

The following theorem [14] provides the prediction interval of the stochastic function value $g(\boldsymbol{x}, \boldsymbol{\xi})$ which can be calculated from its samples.

Theorem 1 *Let* $g(\boldsymbol{x}, \boldsymbol{\xi}^n),\ n = 1, \cdots, N$ *be a set of samples and* $N \geq N_{\min}$. *From a given probability* $\alpha \in (0, 1)$, *the minimum sample size is*

$$N_{\min} = \left\lfloor \frac{1}{\alpha} + 1 \right\rfloor. \tag{9}$$

From the probability $\alpha \in (0, 1)$ *and the sample size* $N > N_{\min}$, *an coefficient* κ *is defined as*

$$\kappa = \sqrt{\frac{N^2 - 1}{N\,(\alpha\, N - 1)}}. \tag{10}$$

By using $\overline{g}(\boldsymbol{x})$ *in (6),* $s^2(\boldsymbol{x})$ *in (7), and* κ *in (10), the prediction interval of* $g(\boldsymbol{x}, \boldsymbol{\xi})$ *is estimated as*

$$\left[\begin{array}{l} \Pr([g^L(\boldsymbol{x}),\ g^U(\boldsymbol{x})] \ni g(\boldsymbol{x}, \boldsymbol{\xi})) \geq 1 - \alpha, \\ g^L(\boldsymbol{x}) = \overline{g}(\boldsymbol{x}) - \kappa\, s(\boldsymbol{x}), \\ g^U(\boldsymbol{x}) = \overline{g}(\boldsymbol{x}) + \kappa\, s(\boldsymbol{x}). \end{array}\right. \tag{11}$$

Proof: From the upper-bound of the floor function, i.e. its argument, the right-hand side of Chebyshev inequality in (8) is simplified as

$$\begin{aligned} \Pr\left(|g(\boldsymbol{x}, \boldsymbol{\xi}) - \overline{g}(\boldsymbol{x})| \geq \lambda \sqrt{\frac{N+1}{N}}\, s(\boldsymbol{x})\right) \\ \leq \frac{N-1+\lambda^2}{N\,\lambda^2}. \end{aligned} \tag{12}$$

Let define a coefficient κ temporarily as

$$\kappa = \lambda \sqrt{\frac{N+1}{N}}. \tag{13}$$

Substituting κ in (13) to (12), we have

$$\begin{aligned} \Pr(|g(\boldsymbol{x}, \boldsymbol{\xi}) - \overline{g}(\boldsymbol{x})| \geq \kappa\, s(\boldsymbol{x})) \\ \leq \frac{N^2 - 1 + N\,\kappa^2}{N^2\,\kappa^2}. \end{aligned} \tag{14}$$

Let describe the probability α as

$$\alpha = \frac{N^2 - 1 + N\,\kappa^2}{N^2\,\kappa^2}. \tag{15}$$

From (14) and (15), we can derive the prediction interval of $g(\boldsymbol{x}, \boldsymbol{\xi})$ as shown in (11). From (15), we can redefine κ by using α as shown in (10). From (13), $\kappa > 1$ holds. From (10) and $\kappa > 1$, the minimum sample size $N_{\min}$ is given by (9). □

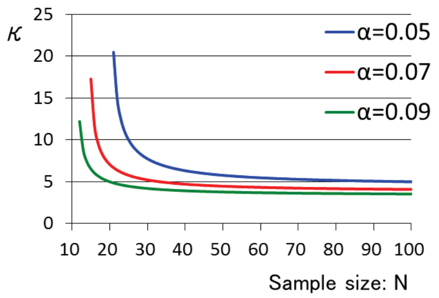


Figure 1: Change of κ in (10) for N and α

It boils down to this. By using (6), (7), and (10), the lower-bound $g^L(\boldsymbol{x})$ of $g(\boldsymbol{x},\,\boldsymbol{\xi})$ shown in (11) is calculated from samples $g(\boldsymbol{x},\,\boldsymbol{\xi}^n)$, $n = 1, \cdots, N$.

Fig. 1 shows the value of coefficient κ in (10) that depends on both the probability α and the sample size N [14]. A limiting value of κ is

$$\lim_{N\to\infty} \kappa = \lim_{N\to\infty} \sqrt{\frac{N^2-1}{N\,(\alpha\,N-1)}} = \sqrt{\frac{1}{\alpha}}. \tag{16}$$

3.2 Problem Formulation

As stated above, CPOP in (3) is transformed into LPOP. By using the lower-bound $g^L(\boldsymbol{x})$ of $g(\boldsymbol{x},\,\boldsymbol{\xi})$ provided by Theorem 1, LPOP is formulated as

$$\left[\begin{array}{ll} \max\limits_{\boldsymbol{x}\in\boldsymbol{X}} & g^L(\boldsymbol{x}) = \overline{g}(\boldsymbol{x}) - \kappa\, s(\boldsymbol{x}) \\ \text{sub. to} & x_1 + x_2 + \cdots + x_D = 1 \end{array}\right. \tag{17}$$

where we suppose that the sample size N used to calculate $g^L(\boldsymbol{x})$ is larger enough than $N_{\min}$ in (9).

Theorem 2 *If $\boldsymbol{x} \in \boldsymbol{X}$ is a feasible solution of LPOP in (17) then $\boldsymbol{x} \in \boldsymbol{X}$ is also a feasible solution of CPOP in (3) under a condition $\gamma = g^L(\boldsymbol{x})$.*

Proof: Assume that $\boldsymbol{x} \in \boldsymbol{X}$ is a feasible solution of LPOP in (17). Therefore, $\boldsymbol{x} \in \boldsymbol{X}$ provides the lower-bound as shown in (11). Since $g^L(\boldsymbol{x}) = \gamma$ holds,

$$\begin{array}{l} \Pr(g(\boldsymbol{x},\,\boldsymbol{\xi}) \ge \gamma) = \Pr([\gamma,\,\infty) \ni g(\boldsymbol{x},\,\boldsymbol{\xi})) \\ \ge \Pr([g^L(\boldsymbol{x}),\,g^U(\boldsymbol{x})] \ni g(\boldsymbol{x},\,\boldsymbol{\xi})) \ge 1-\alpha. \end{array} \tag{18}$$

From (18), the feasible solution of LPOP $\boldsymbol{x} \in \boldsymbol{X}$ satisfies the chance constraint of CPOP in (3). □

3.3 Generation of Samples

We need samples $\boldsymbol{\xi}^n \sim \mathcal{N}(\boldsymbol{\mu},\,\boldsymbol{V})$ $n = 1, \cdots, N$ to calculate the lower-bound $g^L(\boldsymbol{x})$ of $g(\boldsymbol{x},\,\boldsymbol{\xi})$ for LPOP. Therefore, we explain how to generate them. From Cholesky decomposition [15], the covariance matrix $\boldsymbol{V}$ in (2) is decomposed into the form

$$\boldsymbol{V} = \boldsymbol{B}\,\boldsymbol{B}^T \tag{19}$$

where $\boldsymbol{B}$ is a lower triangular matrix:

$$\boldsymbol{B} = \begin{bmatrix} b_{11} & 0 & \cdots & 0 \\ b_{21} & b_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ b_{D1} & b_{D2} & \cdots & b_{DD} \end{bmatrix}. \tag{20}$$

From $\boldsymbol{B}$ in (20) and $\boldsymbol{\mu} = (\mu_1, \cdots, \mu_D)$ in (1), each sample $\boldsymbol{\xi}^n = (\xi_1^n, \cdots, \xi_D^n)$ is generated as

$$\left[\begin{array}{l} \xi_1^n = \mu_1 + b_{11}\,\varepsilon_1^n \\ \xi_2^n = \mu_2 + b_{21}\,\varepsilon_1^n + b_{22}\,\varepsilon_2^n \\ \quad\vdots \\ \xi_D^n = \mu_D + b_{D1}\,\varepsilon_1^n + \cdots + b_{DD}\,\varepsilon_D^n \end{array}\right. \tag{21}$$

where $\varepsilon_i^n \sim \mathcal{N}(0,\,1)$, $i = 1, \cdots, D$ are mutually independent random variables following the standard normal distribution.

4 Optimization Algorithm

In order to obtain the solution of LPOP defined by (17), we employ a basic DE [11]. That is because DE is arguably one of the most powerful stochastic real-parameter optimization algorithms in current use [16]. We describe the procedure of DE for LPOP.

4.1 Initialization

DE holds N_P individuals in the population $\boldsymbol{P}$. The k-th individual $\boldsymbol{v}_k \in \boldsymbol{P} \subseteq [0,\,1]^D$, $k = 1, \cdots, N_P$ is a D-dimensional real vector represented as

$$\boldsymbol{v}_k = (v_{1,k}, \cdots, v_{i,k}, \cdots, v_{D,k}) \tag{22}$$

where $v_{i,k} \in \Re$ and $0 \le v_{i,k} \le 1$, $i = 1, \cdots, D$.

The initial population $\boldsymbol{P}$ is generated randomly. Then each of the individuals $\boldsymbol{v}_k \in \boldsymbol{P}$ is transformed in a solution $\boldsymbol{x} = (x_1, \cdots, x_D) \in \boldsymbol{X}$ of LPOP as

$$x_i = \frac{v_{i,k}}{v_{1,k} + \cdots + v_{D,k}},\ i = 1, \cdots, D. \tag{23}$$

Consequently, the equality constraint in (17), namely $x_1 + \cdots + x_D = 1$, is always satisfied.

The solution $\boldsymbol{x} \in \boldsymbol{X}$ is evaluated N times such as $g(\boldsymbol{x},\,\boldsymbol{\xi}^n)$, $n = 1, \cdots, N$. From the set of samples, the objective function value of LPOP, namely the lower-bound $g^L(\boldsymbol{x})$, is calculated as stated above.

4.2 Strategy of DE

In order to generate a candidate for a new $\boldsymbol{v}_k \in \boldsymbol{P}$, DE uses a basic strategy named "DE/rand/1/bin" [11]. The performance of the strategy usually depends on the values of the control parameters, namely the scale factor $S_F \in (0, 1]$ and the crossover rate $C_R \in [0, 1]$, which are given by users in advance. Due to avoid a strong dependency on the control parameters, a self-adaptive setting of them is employed. The self-adaptive setting is detailed in the literature [17].

Each of the individuals $\boldsymbol{v}_k \in \boldsymbol{P}$ is assigned to "target vector" in turn. Except for the target vector, three other distinct individuals, say $\boldsymbol{v}_{k1}$, $\boldsymbol{v}_{k2}$, and $\boldsymbol{v}_{k3}$, $k \neq k1 \neq k2 \neq k3$, are selected randomly from $\boldsymbol{P}$. From the three individuals, the differential mutation generates a new vector $\boldsymbol{z} = (z_1, \cdots, z_D) \in \Re^D$ called "mutated vector" as

$$\boldsymbol{z} = \boldsymbol{v}_{k1} + S_F\,(\boldsymbol{v}_{k2} - \boldsymbol{v}_{k3}). \tag{24}$$

The binomial crossover [11] between the mutated vector $\boldsymbol{z} \in \Re^D$ in (24) and the target vector $\boldsymbol{v}_k \in \boldsymbol{P}$ generates a candidate for a new individual $\boldsymbol{u} \in \Re^D$ called "trial vector". Each component $u_i \in \Re$ of the trial vector $\boldsymbol{u} = (u_1, \cdots, u_D) \in \Re^D$ is inherited from either $\boldsymbol{z} \in \Re^D$ or $\boldsymbol{v}_k \in \boldsymbol{P}$ as

$$u_i = \begin{cases} z_i; & \text{if } \mathrm{rand}_i \leq C_R \,\vee\, i = i_r \\ v_{i,k}; & \text{otherwise} \end{cases} \tag{25}$$

where $\mathrm{rand}_i \in [0, 1]$ denotes a uniformly distributed random value. The subscript $i_r \in [1, D]$ is also selected randomly, which ensures that the newborn $\boldsymbol{u} \in \Re^D$ differs from the existing $\boldsymbol{v}_k \in \boldsymbol{P}$.

If a component $u_i \in \Re$ of the trial vector $\boldsymbol{u} \in \Re^D$ is made out of the range $[0, 1]$ as a result of the above strategy, it is returned to the range $[0, 1]$ as

$$u_i = \begin{cases} v_{i,k1} + \mathrm{rand}_i\,(0 - v_{i,k1}); & \text{if } u_i < 0 \\ v_{i,k1} + \mathrm{rand}_i\,(1 - v_{i,k1}); & \text{if } u_i > 1 \end{cases} \tag{26}$$

where $\boldsymbol{v}_{k1} \in \boldsymbol{P}$ has been used in (24) too [11].

4.3 Survival Selection

The trial vector $\boldsymbol{u} \in \Re^D$ is transformed into a solution of LPOP as shown in (23). The new solution is also evaluated N times to get the samples of its function value. Then the objective function value is calculated for the solution. The trial vector $\boldsymbol{u} \in \Re^D$ is compared with the target vector $\boldsymbol{v}_k \in \boldsymbol{P}$ in their objective function values. As a result, if $\boldsymbol{u} \in \Re^D$ is not worse than $\boldsymbol{v}_k \in \boldsymbol{P}$, $\boldsymbol{v}_k \in \boldsymbol{P}$ is replaced by $\boldsymbol{u} \in \Re^D$ immediately. Otherwise, the trial vector $\boldsymbol{u} \in \Re^D$ is discarded.

Algorithm 1 : DE for LPOP

$\boldsymbol{P} :=$ GENERATE_INITIAL_POPULATION(N_P);
$\forall \boldsymbol{v}_k \in \boldsymbol{P}$: Evaluate $g^L(\tau(\boldsymbol{v}_k))$;
repeat
  for $k := 1$ to N_P **do**
    $\boldsymbol{u} :=$ STRATEGY$(\boldsymbol{v}_k, \boldsymbol{P})$;
    Evaluate $g^L(\tau(\boldsymbol{u}))$;
    if $g^L(\tau(\boldsymbol{u})) \geq g^L(\tau(\boldsymbol{v}_k))$ **then**
      $\boldsymbol{v}_k := \boldsymbol{u}$; /* $\boldsymbol{v}_k \in \boldsymbol{P}$ */
    end if
  end for
until a termination condition is satisfied;
$\boldsymbol{x}^\star :=$ SELECT_BEST_FEASIBLE$(\boldsymbol{P})$;

Algorithm 1 provides the pseudo-code of DE for LPOP. The function $\tau : [0, 1]^D \to \boldsymbol{X}$ means that an individual $\boldsymbol{v}_k \in [0, 1]^D$ is transformed into a solution $\boldsymbol{x} \in \boldsymbol{X}$ of LPOP as shown in (23). The best feasible solution $\boldsymbol{x}^\star \in \boldsymbol{X}$ of LPOP is returned in the end.

5 Numerical Experiment

5.1 Example of CPOP

We consider an example of CPOP in which the ratios of four assets x_i, $i = 1, \cdots, 4$ are optimized under a probability $\alpha \in (0, 1)$. The mean vector $\boldsymbol{\mu} \in \Re^4$ and the covariance matrix $\boldsymbol{V}$ in (1) are given as

$$\boldsymbol{\mu} = (0.05,\, 0.10,\, 0.15, 0.20),$$

$$\boldsymbol{V} = \begin{bmatrix} 0.0004 & -0.0006 & 0.0001 & -0.0006 \\ -0.0006 & 0.0016 & -0.0012 & 0.0006 \\ 0.0001 & -0.0012 & 0.0036 & -0.0014 \\ -0.0006 & 0.0006 & -0.0014 & 0.0064 \end{bmatrix}.$$

5.2 Example of LPOP

We transform the above CPOP into LPOP. From Fig. 1, the sample size is chosen as $N = 80$. As a result of the Cholesky decomposition of $\boldsymbol{V}$, the lower triangle matrix $\boldsymbol{B}$ in (20) is obtained as

$$\boldsymbol{B} = \begin{bmatrix} 0.0200 & 0.0000 & 0.0000 & 0.0000 \\ -0.0280 & 0.0286 & 0.0000 & 0.0000 \\ 0.0060 & -0.0361 & 0.0475 & 0.0000 \\ -0.0320 & -0.0090 & -0.0331 & 0.0648 \end{bmatrix}.$$

5.3 Experimental Setup

The proposed DE for LPOP was coded by the Java language. As the termination condition of DE, the maximum number of generation was fixed to 200. The population size was chosen as $N_P = 20$. Thereby, we applied DE to an instance of LPOP 30 times.

Table 1: Objective function value of LPOP

α	MAX	AVE	VAR
0.05	0.0681	0.0652	0.0011
0.10	0.0859	0.0817	0.0019
0.15	0.1034	0.0962	0.0023
0.20	0.1153	0.1071	0.0025
0.25	0.1193	0.1145	0.0025
0.30	0.1262	0.1212	0.0023

Table 2: The best solution of LPOP

α	x_1	x_2	x_3	x_4
0.05	0.4449	0.2587	0.1791	0.1171
0.10	0.1782	0.2909	0.3319	0.1988
0.15	0.0364	0.2712	0.3664	0.3257
0.20	0.0043	0.1482	0.4388	0.4085
0.25	0.0019	0.0990	0.5058	0.3930
0.30	0.0070	0.1321	0.4189	0.4419

5.4 Experimental Result

Several risks $\alpha \in (0,\ 1)$ were given for the example of CPOP. As stated above, each CPOP was transformed into an instance of LPOP. Table 1 shows the result of experiment. The maximum (MAX), average (AVE), and variance (VAR) of the objective function values of LPOP achieved by the solutions obtained by DE are shown in Table 1. From the results in Table 1, we can confirm the robustness of DE for LPOP.

Fig. 2 shows the change of the total return γ of CPOP for the risk α which is provided by the best solutions of LPOP shown in Table 1. From Fig. 2, it is observed that the total return γ of CPOP increases in proportion to the amount of risk α.

Table 2 shows the best solution $\boldsymbol{x}^{\star} \in \boldsymbol{P}$ obtained by DE for the respective α values. From the result in Table 2, the best solution depends on the risk α. Fig. 3 illustrates the ratios of assets composing the best solutions of LPOP shown in Table 2. The values of x_1 and x_2 dominate the investment when the risk is small: $\alpha = 0.05$. On the other hand, the values of x_3 and x_4 dominate the investment when the risk is large: $\alpha = 0.25$. Therefore, the former assets might be safety and the latter assets might be risky.

6 Conclusion

A new approach for solving CPOP without using the Monte Carlo simulation was proposed. By using the lower-bound of stochastic function value, which were derived from Chebyshev inequality [13], CPOP was transformed into LPOP. Besides, it was proved that the feasible solution of LPOP is also feasible for CPOP.

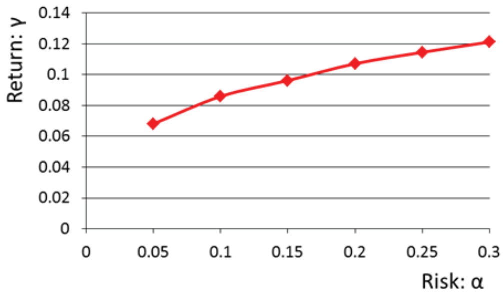
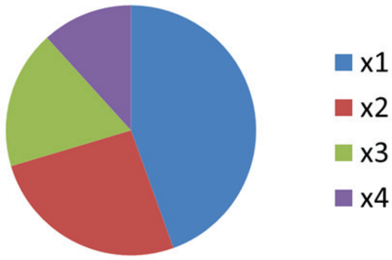
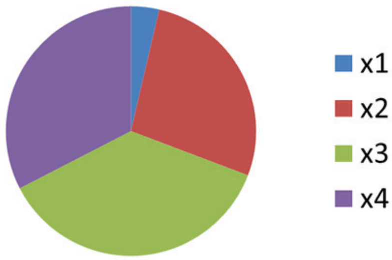


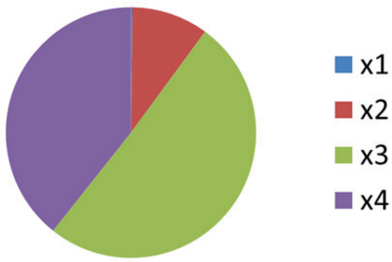
Figure 2: Change of return γ for risk α



(a) Risk: $\alpha = 0.05$



(b) Risk: $\alpha = 0.15$



(c) Risk: $\alpha = 0.25$

Figure 3: The best solution $\boldsymbol{x}^{\star} \in \Re^4$ for risk α

Then, in order to solve LPOP, a basic DE was used. Through the numerical experiment on a test problem, we could confirm that the total return of CPOP was increased in proportion to the amount of risk α.

In the future work, we reduce the computational burden of the proposed approach. For solving LPOP with the basic DE, every individual has to be evaluated N times. Compared with the primitive Monte Carlo simulation, the sample size N is relatively small. However, the multiple sampling of every individual is still expensive. Therefore, we need to introduce some sample-saving techniques [10, 14] into the basic DE. Furthermore, we would like to evaluate the proposed approach on practical portfolio optimization problems which contain a large number of assets [18].

Acknowledgements: The research was supported by JSPS KAKENHI Grant Number 17K06508.

References:

[1] H. Markowitz: Portfolio selection, *Journal of Finance*, 7, 1952, pp. 77–91.

[2] S. Kataoka: A stochastic programming model, *Econometrica*, 31, 1963, pp. 181–196.

[3] H. Chang and W. Wei: Portfolio optimization with random liability in the stochastic interest rate environments, *WSEAS Trans. on Systems and Control*, 10, 2015, pp. 752–762.

[4] B. K. Pagnoncelli, S. Ahmed, and A. Shapiro: Sample average approximation method for chance constrained programming: theory and applications, *Journal of Optimization Theory and Applications*, 142, 2009, pp. 399–416.

[5] H. Tanaka, P. Guo, and I. B. Türksen: Portfolio selection based on fuzzy probabilities and possibility distributions, *Fuzzy Set and Systems*, 111, 2000, pp. 387–397.

[6] L. Yan: Chance-constrained programming model for portfolio selection in uncertain environment, *Modern Applied Science*, 3(10), 2009, pp. 89–95.

[7] W. Chen and S. Tan: Interval random dependent-chance programming and its application to portfolio selection, *Proc. of IEEE International Conference of Fuzzy Systems*, 2009, pp. 626–630.

[8] M. M. Mehlawat and P. Gupta: Fuzzy chance-constrained multiobjective portfolio selection model, *IEEE Trans. on Fuzzy Systems*, 22(3), 2014, pp. 653–671.

[9] Z. Qin: *Uncertain Portfolio Optimization*, Springer, 2016.

[10] K. Tagawa: Chebyshev inequality based approach to chance constrained optimization problems using differential evolution, *Advances in Swarm Intelligence*, LNCS 10385, Springer, 2017, pp. 440–448.

[11] K. V. Price, R. M. Storn, and J. A. Lampinen: *Differential Evolution - A Practical Approach to Global Optimization*, Springer, 2005.

[12] P. Tchébychef: Des valeurs moyennes, *Journal de Mathématiques Pures et Appliquées*, 2(12), 1867, pp. 177–184.

[13] J. G. Saw, M. C. K. Yang, and T. C. Mo: Chebyshev inequality with estimated mean and variance, *The American Statistician*, 38(2), 1984, pp. 130–132.

[14] K. Tagawa: Worst case optimization using Chebyshev inequality, *Bioinspired Optimization Methods and their Applications*, Jožef Stefan Institute, 2016, pp. 173–185.

[15] J. E. Gentle: *Numerical Linear Algebra for Applications in Statistics*, Springer, 1998.

[16] S. Das, S. S. Mullick, and P. N. Suganthan: Recent advances in differential evolution - an updated survey, *Swarm and Evolutionary Computation*, 27, 2016, pp. 1–30.

[17] J. Brest, S. Greiner, B. Bošković, M. Merink, and V. Žumer: Self-adapting control parameters in differential evolution: a comparative study on numerical benchmark problems, *IEEE Trans. on Evolutionary Computation*, 10(6), 2006, pp. 646–657.

[18] M. Hirschberger, Y. Qi, and R. E. Steuer: Randomly generating portfolio-selection covariance matrices with specified distributional characteristics, *European Journal of Operational Research*, 117, 2007, pp. 1610–1625.