GPIF Working Papers

The use of one-period optimization models for long-term investments: In terms of rebalancing frequency and sampling frequency

GPIF

November 7, 2024
Koichi Miyazaki
Takuya Kiriu
Keiji Watanabe

EXECUTIVE SUMMARY

- In long-term investments such as public pension funds, it is necessary to pay attention to the probability of the amount of assets falling below the required level at a future point in time (downside probability). In actual investment, a policy asset mix should be established and then rebalanced in a timely and appropriate manner. If a one-period optimization model based on a mean-variance model is used to formulate the policy asset mix, it is required to confirm the downside probability when the policy asset mix derived from the model is managed faithfully.
- This study shows that the condition that "the frequencies of data sampling and rebalancing assumed in the operation are both sufficiently high" should be satisfied in order to explicitly grasp the correspondence between the parameters of the one-period optimization model such as mean and variance and the downside probability from the normal distribution, and examine the use of the one-period optimization model in long-term investment from the perspective of the rebalancing and sampling frequencies.

(Note: This working paper is a compilation of research results by GPIF staff, and the contents and opinions expressed in the text do not represent the official views of the GPIF. We would also like to thank Mr. Hikaru Takeuchi, a GPIF staff member, for his helpful comments in preparing this working paper. We hereby note and thank him for his comments.

1. Introduction

When constructing portfolios, one-period optimization models are mainly employed due to the ease of model construction and solution. In GPIF, assets are managed within the allowed range to avoid significant deviation from the policy asset mix (see GPIF (2023) for detail). Therefore, investment decisions, such as major asset allocation changes, are made at the time the policy asset mix is determined. Based on the assumption of such investment process, the policy asset mix is constructed using a one-period optimization model. However, since it is assumed that public pension funds are expected to use the invested assets for pension payments in the future, it is necessary to measure the probability that the assumed amount of assets will fall below the required level at a future point in time (downside probability) when the actual portfolio is faithfully managed according to the policy asset mix determined using the one-period optimization model.

The most prominent one-period model is the mean-variance model proposed by Markowitz (1952). Honda (2019) explains that the mean-variance model still provides a very effective methodology for analyzing financial markets and constructing portfolio strategies. Honda (2019) discusses the issue closely related to the sampling frequency, which is the subject of this study. One of the topics that Honda (2019) addresses is the issue of estimating parameters used in mean and variance models, and points out some caveats when using parameters, based on previous studies such as Levy and Roll (2010) and Kan and Zhou (2007). Specifically, "Even if the estimated parameters are unbiased estimators (i.e., even if the number of assets that are estimated to be larger than their true values and the number of assets that are estimated to be smaller are roughly equal), when obtained estimates are used as input values for the mean and variance model to optimize portfolios, the derived efficient frontier tends to expand." In light of this point, when attempting optimization based on estimated parameters, it is desirable that the confidence interval of the estimated parameter be as small as possible, in addition to being unbiased. Therefore, when estimating parameters from historical data, the sampling frequency should be sufficient to ensure a sufficient number of return data.

In order to ensure the actual portfolio does not deviate significantly from the policy asset mix in long-term fund management, it is necessary to appropriately rebalance the portfolio as needed when the weight of the managed portfolio deviates from that of the policy asset mix. Shimizu and Uchiyama (2017) are among the preceding studies that analyzed the impact of systematic rebalancing on performance (rebalancing premium). After comparing the rebalancing premium from a Monte Carlo simulation using parameters estimated from the market assuming a geometric Brownian motion model and that obtained from rebalancing actual market data, they found that the contribution of rebalancing to performance improvement is limited when "investment opportunities do not change (estimated parameters are fixed)." In addition, if the rebalancing premium from the actual rebalancing is significantly larger, the study suggests that it is due to the impact of "fluctuations in investment opportunities," such as "risk reversal" (a phenomenon in which an asset with a low past performance subsequently performs well over time, while an asset with a high past performance subsequently performs poorly).

The purpose of this study is to show that the condition that "the frequency of data sampling and the frequency of rebalancing are both sufficiently high" should be satisfied in order to justify the relation between the parameters of the one-period optimization model with mean and variance and the downside probability from the normal distribution, and to examine the points to be considered when using the one-period optimization model for long-term investment in terms of the rebalancing frequency and sampling frequency.

The structure of this paper is as follows. In Section 2, we adopt the geometric Brownian motion model as in Shimizu and Uchiyama (2017) and organize our ideas on rebalancing frequency (continuous rebalancing and buy and hold) in line with the purpose of this study. In Section 3, we discretize the geometric Brownian motion to more explicitly capture the effects of rebalancing frequency and sampling frequency (of return data used for parameter estimation) on downside probabilities. In Section 4, based on numerical examples, we review the differences in the size of the confidence intervals for the parameters due to the differences in sampling frequencies. In the final section, a summary and future issues are added.

2. The case where each asset follows a geometric Brownian motion

2.1 Geometric Brownian motion and rebalancing frequency

The price at the time t for asset i (i = 1,2,3,4), S_t^i, follows the geometric Brownian motion in equation

(1).

$$\frac{dS_t^i}{S_t^i} = \mu^i dt + \sigma^i dB_t^i \tag{1}$$

Where, B_t^i is the standard Brownian motion, and $dB_t^i \cdot dB_t^j dt = \rho_{i,j} dt$. For details of Brownian motion, see Morimura and Kijima (1991).

To handle each weight in the portfolio, the weight vector of the four assets (four assets are assumed because the current GPIF policy asset mix consists of four asset classes, but the same argument holds for the general case) $\boldsymbol{w} = (w_1, w_2, w_3, w_4)$ and drift vector $\boldsymbol{\mu} = (\mu_1, \mu_2, \mu_3, \mu_4)$ and diffusion matrices $\boldsymbol{\Omega}$ are introduced.

(Case of continuous rebalancing)

Based on equation (1), the stochastic process followed by the portfolio return with weight $\boldsymbol{w}$ is expressed as equation (2).

$$\begin{aligned}\frac{dS_t^P}{S_t^P} &= \sum_{i=1}^{4} w_i \frac{dS_t^i}{S_t^i} = \sum_{i=1}^{4} w_i \left(\mu^i dt + \sigma^i dB_t^i\right) \\ &= \left(\sum_{i=1}^{4} w_i \mu^i\right) dt + \sum_{i=1}^{4} w_i \sigma^i dB_t^i = \boldsymbol{w\mu}^T dt + \sqrt{\boldsymbol{w\Omega w}^T} dB_t\end{aligned} \tag{2}$$

Equation (2) implies that at every instant of time t the distribution of the portfolio return at the next smallest time interval dt follows $N(\boldsymbol{w\mu}^T dt,\ \boldsymbol{w\Omega w}^T dt)$. This also means that the portfolio weights are always kept at $\boldsymbol{w}$, in other words, the portfolio is continuously rebalanced.

If equation (2) is always satisfied (i.e., continuous rebalancing), the total asset value S_0^P at time 0 is grown up to S_t^P by integrating equation (2) from time point 0 to time point t.

$$S_t^P = S_0^P e^{\left(\boldsymbol{w\mu}^T - \frac{1}{2}\boldsymbol{w\Omega w}^T\right)t + \sqrt{\boldsymbol{w\Omega w}^T} B_t} \tag{3}$$

Using equation (3), we obtain the portfolio gross return from time point 0 to time point t, $R_{0\to t}^P$ as the stochastic process expressed by equation (4).

$$R_{0\to t}^P = \frac{S_t^P}{S_0^P} = e^{\left(\boldsymbol{w\mu}^T - \frac{1}{2}\boldsymbol{w\Omega w}^T\right)t + \sqrt{\boldsymbol{w\Omega w}^T} B_t} \tag{4}$$

From equation (4), the expected value and the variance of the portfolio gross return $R_{0\to t}^P$ are given by equations (5) and (6), respectively.

$$\begin{aligned}E(R_{0\to t}^P) &= E\left(e^{\left(\boldsymbol{w\mu}^T - \frac{1}{2}\boldsymbol{w\Omega w}^T\right)t + \sqrt{\boldsymbol{w\Omega w}^T} B_t}\right) = e^{\left(\boldsymbol{w\mu}^T - \frac{1}{2}\boldsymbol{w\Omega w}^T\right)t} E\left(e^{\sqrt{\boldsymbol{w\Omega w}^T} B_t}\right) \\ &= e^{\left(\boldsymbol{w\mu}^T - \frac{1}{2}\boldsymbol{w\Omega w}^T\right)t} e^{\left(\frac{1}{2}\boldsymbol{w\Omega w}^T\right)t} = e^{(\boldsymbol{w\mu}^T)t}\end{aligned} \tag{5}$$

$$V(R_{0\to t}^P) = V\left(e^{\left(\boldsymbol{w\mu}^T - \frac{1}{2}\boldsymbol{w\Omega w}^T\right)t + \sqrt{\boldsymbol{w\Omega w}^T}B_t}\right) = e^{2\left(\boldsymbol{w\mu}^T - \frac{1}{2}\boldsymbol{w\Omega w}^T\right)t} V\left(e^{\sqrt{\boldsymbol{w\Omega w}^T}B_t}\right)$$
$$= e^{2\left(\boldsymbol{w\mu}^T - \frac{1}{2}\boldsymbol{w\Omega w}^T\right)t}\left(E\left(\left(e^{\sqrt{\boldsymbol{w\Omega w}^T}B_t}\right)^2\right) - \left(E\left(e^{\sqrt{\boldsymbol{w\Omega w}^T}B_t}\right)\right)^2\right)$$
$$= e^{2\left(\boldsymbol{w\mu}^T - \frac{1}{2}\boldsymbol{w\Omega w}^T\right)t}\left(E\left(e^{2\sqrt{\boldsymbol{w\Omega w}^T}B_t}\right) - \left(E\left(e^{\sqrt{\boldsymbol{w\Omega w}^T}B_t}\right)\right)^2\right) \qquad (6)$$
$$= e^{2\left(\boldsymbol{w\mu}^T - \frac{1}{2}\boldsymbol{w\Omega w}^T\right)t}\left(e^{2\left(\boldsymbol{w\mu}^T - \frac{1}{2}\boldsymbol{w\Omega w}^T\right)t} - \left(E\left(e^{\sqrt{\boldsymbol{w\Omega w}^T}B_t}\right)\right)^2\right)$$
$$= e^{2\boldsymbol{w\mu}^T t}\left(e^{\boldsymbol{w\Omega w}^T t} - 1\right)$$

From equation (4), the expected value and the variance of the portfolio log return $ln(R_{0\to t}^P)$ are attained as equations (7) and (8), respectively.

$$E\left(ln(R_{0\to t}^P)\right) = E\left(\left(\boldsymbol{w\mu}^T - \frac{1}{2}\boldsymbol{w\Omega w}^T\right)t + \sqrt{\boldsymbol{w\Omega w}^T}B_t\right) = \left(\boldsymbol{w\mu}^T - \frac{1}{2}\boldsymbol{w\Omega w}^T\right)t \qquad (7)$$

$$V\left(ln(R_{0\to t}^P)\right) = V\left(\left(\boldsymbol{w\mu}^T - \frac{1}{2}\boldsymbol{w\Omega w}^T\right)t + \sqrt{\boldsymbol{w\Omega w}^T}B_t\right) = \boldsymbol{w\Omega w}^T t \qquad (8)$$

Organize the discussion so far. The price at the time t for asset i (i = 1,2,3,4), S_t^i, follows the geometric Brownian motion in equation (1), and the portfolio weight is always $\boldsymbol{w}$ with continuously rebalanced from time point 0 to time point t, the initial portfolio value of 1 will be $R_{0\to t}^P$ at time point t. The expected value of the portfolio amount is $E(R_{0\to t}^P) = e^{(\boldsymbol{w\mu}^T)t}$ and the variance of the portfolio amount is $V(R_{0\to t}^P) = e^{2\boldsymbol{w\mu}^T t}\left(e^{\boldsymbol{w\Omega w}^T t} - 1\right)$. Also, the probability distribution of the portfolio return $R_{0\to t}^P = e^X$ follows a lognormal distribution with the stochastic variable $X \sim N\left(\left(\boldsymbol{w\mu}^T - \frac{1}{2}\boldsymbol{w\Omega w}^T\right)t, (\boldsymbol{w\Omega w}^T)t\right)$. At time point t, the probability that the amount of the portfolio $R_{0\to t}^P$ is below the threshold K is given by $P(R_{0\to t}^P \leq K) = P\left(ln(R_{0\to t}^P) \leq ln(K)\right) = P\left(X \leq ln(K)\right)$. Thus, the downside probability is obtained exactly (without using any approximation).

(Case without rebalancing)

Next, we consider the portfolio gross return when no rebalancing is performed from time point 0 to time point t. The amount of assets at time 0 for asset i (i = 1,2,3,4), S_0^i, will be the amount of assets at time point t, S_t^i, in equation (9) by integrating equation (1) from time point 0 to time point t.

$$S_t^i = S_0^i e^{\left(\mu^i - \frac{1}{2}{\sigma^i}^2\right)t + \sigma^i B_t^i} \qquad (9)$$

The asset value of the portfolio at time point 0, S_0^P, grows up to the asset value of the portfolio at time point t, $S_t^{P,N}$, in equation (10).

$$S_t^{P,N} = \sum_{i=1}^{4} S_0^i e^{\left(\mu^i - \frac{1}{2}{\sigma^i}^2\right)t + \sigma^i B_t^i} \qquad (10)$$

Where, $S_0^i = w_i S_0^P$.

Using equation (10), the stochastic variable $R_{0\to t}^{P,N}$ representing the portfolio gross return from time point 0 to time point t is given by equation (11).

$$R_{0\to t}^{P,N} = \frac{S_t^{P,N}}{S_0^P} = \sum_{i=1}^{4} w_i e^{\left(\mu^i - \frac{1}{2}\sigma^{i^2}\right)t + \sigma^i B_t^i} \tag{11}$$

From equation (11), the expected value and the variance of the portfolio gross return $R_{0\to t}^{P,N}$ are equations (12) and (13), respectively.

$$\begin{aligned} E\left(R_{0\to t}^{P,N}\right) &= E\left(\sum_{i=1}^{4} w_i e^{\left(\mu^i - \frac{1}{2}\sigma^{i^2}\right)t + \sigma^i B_t^i}\right) = \sum_{i=1}^{4} w_i E\left(e^{\left(\mu^i - \frac{1}{2}\sigma^{i^2}\right)t + \sigma^i B_t^i}\right) \\ &= \sum_{i=1}^{4} w_i e^{\mu^i t} \end{aligned} \tag{12}$$

$$\begin{aligned} V\left(R_{0\to t}^{P,N}\right) &= V\left(\sum_{i=1}^{4} w_i e^{\left(\mu^i - \frac{1}{2}\sigma^{i^2}\right)t + \sigma^i B_t^i}\right) \\ &= \sum_{i=1}^{4} w_i^2 e^{2\mu^i t}\left(e^{\sigma^{i^2} t} - 1\right) + 2\sum_{i\neq j} w_i w_j e^{(\mu^i + \mu^j)t}\left(e^{\rho_{i,j}\sigma^i\sigma^j t} - 1\right) \end{aligned} \tag{13}$$

Assuming that the initial amount of the portfolio is 1, if the portfolio is not rebalanced at all from time point 0 to time point t, the downside probability at time point t is formally given by $P\left(R_{0\to t}^{P,N} \leq K\right)$. However, the lognormal distribution does not have the reproductive property appeared in the distribution such as the normal distribution, and the distribution of the portfolio amount at time point t is not explicitly attained. Due to the reason, the computation of the downside probability at time point t is not exact and is a bit laborious. Toward investigation how the downside probability is related to the parameters of the one-period optimization model, assuming that $\mu^i t$ are small and using the first order Maclaurin expansion of ($e^x \cong 1 + x$), the right-hand side of the equation (12) in the expected value is approximated to equation (14).

$$\begin{aligned} E\left(R_{0\to t}^{P,N}\right) &= \sum_{i=1}^{4} w_i e^{\mu^i t} \cong \sum_{i=1}^{4} w_i\left(1 + \mu^i t\right) = \sum_{i=1}^{4} w_i + \sum_{i=1}^{4} w_i \mu^i t \\ &= 1 + \boldsymbol{w}\boldsymbol{\mu}^T t \end{aligned} \tag{14}$$

We attempt first-order approximation of the right-hand side of equation (13) in the variance. Assuming that $e^{2\mu^i t} \cong 1$ and parameters such as $\sigma^i t$ are small and using the first order Maclaurin expansion of ($e^x \cong 1 + x$), the right-hand side of equation (13) in the variance is approximated to equation (15).

$$\begin{aligned} V\left(R_{0\to t}^{P,N}\right) &= \sum_{i=1}^{4} w_i^2 e^{2\mu^i t}\left(e^{\sigma^{i^2} t} - 1\right) + 2\sum_{i\neq j} w_i w_j e^{(\mu^i + \mu^j)t}\left(e^{\rho_{i,j}\sigma^i\sigma^j t} - 1\right) \\ &\cong \sum_{i=1}^{4} w_i^2 \sigma^{i^2} t + 2\sum_{i\neq j} w_i w_j \rho_{i,j}\sigma^i\sigma^j t = (\boldsymbol{w}\boldsymbol{\Omega}\boldsymbol{w}^T)t \end{aligned} \tag{15}$$

In this way, even in the case of no rebalancing, the expected value and the variance of the portfolio amount $R_{0\to t}^{P,N}$ are able to be approximately derived. However, as the investment period t becomes longer, the accuracy of the approximation rapidly deteriorates.

3. 3. A discrete model where the return of each asset follows a normal

distribution

3.1 Discretization of geometric Brownian motion and data sampling frequency

The price at the time of n for each asset's i $(i = 1,2,3,4)$, S_n^i, follows the geometric Brownian motion. Then, the discretized return X_k^i follows the normal distribution in equation (16).[1]

$$X_{k,\Delta t}^i = \frac{S_{k+\Delta t}^i - S_k^i}{S_k^i} = \mu_{\Delta t}^i \Delta t + \sigma_{\Delta t}^i \varepsilon_{k,\Delta t}^i \tag{16}$$

Where, $\varepsilon_{k,\Delta t}^i$ is assumed to follow a time-series-wise independent normal distribution with mean zero and variance Δt. In addition, the correlation structure among assets is assumed to be $\mathrm{Cov}(\varepsilon_{k,\Delta t}^i, \varepsilon_{k,\Delta t}^j) = \rho_{i,j}$ and there is no time-series-wise cross-correlation. When using this discrete model, the approximation accuracy to the continuous model given by equation (1) decreases as the data sampling frequency decreases. The daily $(\Delta t = 1 / 252)$ sampling frequency, we can generate time series data of daily returns from daily sampled asset prices using the left-hand side of equation (16), and then take their expected value and standard deviation to obtain the estimated value $\mu_{1/252}^i$ and $\sigma_{1/252}^i$ respectively. The concept is the same when the sampled data is divided into monthly $(\Delta t = 1 / 12)$, quarterly $(\Delta t = 1 / 4)$, semi-annual $(\Delta t = 1 / 2)$, and annual $(\Delta t = 1)$.

3.2 Frequency of rebalancing and data sampling

In this section, we examine the rebalancing frequency and data sampling frequency using discretized geometric Brownian motion. We begin to compare the process of deriving the downside probability between the case of rebalancing and the case of no rebalancing in semiannual (twice a year) data sampling frequency and rebalancing for one year (the first half year and the second half year), then extend the comparison to the case N annual frequency and rebalancing N times a year.

From the price at the time of n for asset i, $S_{n,1/2}^i$, the returns obtained during the first and second half year are expressed in $X_{1,1/2}^i, X_{2,1/2}^i$, respectively.

---

[1] In this study, in order to examine the impact of rebalancing frequency and sampling frequency on the precise derivation of downside probability at a future point in time, a discretization of the continuous model (where the arithmetic returns follow a normal distribution), as in equation (16), was selected as the discrete model. Another discrete model is one in which the logarithmic returns follow a normal distribution, as in equation (16') (this is nested in Miyazaki (2005)), which assumes that the differences after the Box-Cox transformation follow a normal distribution. While the modeling using equation (16') can describe multi-period returns as the sum of one-period returns, it is not used in this study because it complicates the treatment of portfolio returns (they are not normally or lognormally distributed).

$$X_{k,\Delta t}^i = ln\left(\frac{S_{k+\Delta t}^i}{S_k^i}\right) = \mu_{\Delta t}^i \Delta t + \sigma_{\Delta t}^i \varepsilon_{k,\Delta t}^i \tag{16'}$$

(Rebalancing portfolio according to data sampling frequency)

Using the notation in equation (16) and taking one-unit time interval as six months, one period return is a six-month return, and the six-month return of a portfolio whose initial weighting is $\boldsymbol{w}$ is $\sum_{i=1}^{4} w_i X_{1,1/2}^{i}$. If the current portfolio amount is written as $S_{0,1/2}^{P}$, the amount of the portfolios at the end of six months and one year are $S_{1,1/2}^{P}$ and $S_{2,1/2}^{P}$, respectively. Using the expression that the six-month return of a portfolio whose initial weighting is $\boldsymbol{w}$ is $\sum_{i=1}^{4} w_i X_{1,1/2}^{i}$, the portfolio values at the end of six months and one year $S_{1,1/2}^{P}$ and $S_{2,1/2}^{P}$ are related to the six-month returns $X_{1,1/2}^{i}, X_{2,1/2}^{i}$ as equation (17) and equation (18), respectively.

$$S_{1,1/2}^{P} = S_{0,1/2}^{P}\left(1 + \sum_{i=1}^{4} w_i X_{1,1/2}^{i}\right) \qquad (17)$$

$$S_{2,1/2}^{P} = S_{1,1/2}^{P}\left(1 + \sum_{i=1}^{4} w_i X_{2,1/2}^{i}\right) \qquad (18)$$

Using equations (17) and (18), the stochastic variable representing the portfolio gross return from time point 0 (the current point in time) to time point 2 (one year later) (the amount of money invested in one unit at time point 0 that will be available at time point 2) $R_{0\to 2,1/2}^{P}$ is given by equation (19). (corresponding to equation (4) in the continuous model)

$$\begin{aligned} R_{0\to 2,1/2}^{P} = \frac{S_{2,1/2}^{P}}{S_{0,1/2}^{P}} &= \frac{S_{0,1/2}^{P}\left(1 + \sum_{i=1}^{4} w_i X_{1,1/2}^{i}\right)\left(1 + \sum_{i=1}^{4} w_i X_{2,1/2}^{i}\right)}{S_{0,1/2}^{P}} \\ &= \left(1 + \sum_{i=1}^{4} w_i X_{1,1/2}^{i}\right)\left(1 + \sum_{i=1}^{4} w_i X_{2,1/2}^{i}\right) \end{aligned} \qquad (19)$$

Using the notation in equation (16) and taking one-unit time interval as $1/N$ year, one period return is a 1 $/N$ year return, and is denoted as $X_{1,1/N}^{i}$. The stochastic variable of the portfolio gross return from time point 0 (at the present time) to time point N (how much one unit of money invested at time point 0 is worth at time point N) $R_{0\to N,1/N}^{P}$ is given by equation (20).

$$R_{0\to N,1/N}^{P} = \frac{S_{N,1/N}^{P}}{S_{0,1/N}^{P}} = \prod_{j=1}^{N}\left(1 + \sum_{i=1}^{4} w_i X_{j,1/N}^{i}\right) \qquad (20)$$

At time point N, the probability that the amount of the portfolio $R_{0\to N,1/N}^{P}$ is below the threshold K is given by $P\left(R_{0\to N,1/N}^{P} \le K\right) = P\left(\prod_{j=1}^{N}\left(1 + \sum_{i=1}^{4} w_i X_{j,1/N}^{i}\right) \le K\right) = P\left(log\left(\prod_{j=1}^{N}\left(1 + \sum_{i=1}^{4} w_i X_{j,1/N}^{i}\right)\right) \le logK\right) = P\left(\sum_{j=1}^{N} log\left(1 + \sum_{i=1}^{4} w_i X_{j,1/N}^{i}\right) \le logK\right) \cong P\left(\sum_{j=1}^{N}\sum_{i=1}^{4} w_i X_{j,1/N}^{i} \le logK\right)$. The sampling frequency N is sufficiently large and $\sum_{i=1}^{4} w_i X_{j,1/N}^{i}$ is sufficiently small, approximation $log\left(1 + \sum_{i=1}^{4} w_i X_{j,1/N}^{i}\right) \cong \sum_{i=1}^{4} w_i X_{j,1/N}^{i}$ holds. The probability distribution followed by the stochastic variable $\sum_{i=1}^{4} w_i X_{1,1/N}^{i}$ is the normal distribution with the expected values and variance-covariance matrices estimated by using $1/N$-year return data $\boldsymbol{\mu}_{\mathbf{1/N}} = \left(\mu_{1,1/N}, \mu_{2,1/N}, \mu_{3,1/N}, \mu_{4,1/N}\right)$ and $\boldsymbol{\Omega}_{\mathbf{1/N}}$, respectively. In the rebalancing case, if the sampling frequency N is sufficiently large, the downside probability can be accurately obtained from the normal distribution.

(Without rebalancing portfolio according to data sampling frequency)

Next, consider the portfolio gross return when no rebalancing is performed from time point 0 to time point

2. The price at time point 0 for asset i (i = 1,2,3,4), $S_{0,1/2}^{i}$, becomes $S_{2,1/2}^{i}$ at time point 2 in equation (21).

$$S_{2,1/2}^{i} = S_{0,1/2}^{i}\left(1 + X_{1,1/2}^{i}\right)\left(1 + X_{2,1/2}^{i}\right) \tag{21}$$

Using this, the total asset value S_0^P at time point 0 becomes $S_{2,1/2}^{P,N}$ at time point 2 in equation (22).

$$S_{2,1/2}^{P,N} = S_{0,1/2}^{P}\sum_{i=1}^{4} w_i \left(1 + X_{1,1/2}^{i}\right)\left(1 + X_{2,1/2}^{i}\right) \tag{22}$$

Where, $S_{0,1/2}^{i} = w_i S_{0,1/2}^{P}$.

Using equation (22), the stochastic variable representing the portfolio gross return from time point 0 to time point 2 $R_{0\to 2,1/2}^{P,N}$ is given by equation (23).

$$R_{0\to 2}^{P,N} = \frac{S_2^{P,N}}{S_0^P} = \sum_{i=1}^{4} w_i\left(1 + X_{1,1/2}^{i}\right)\left(1 + X_{2,1/2}^{i}\right) \tag{23}$$

Taking one-unit time interval as $1/N$ year, one period return is a $1/N$ year return, and is denoted as $X_{1,1/N}^{i}$. The stochastic variable of the portfolio gross return from time point 0 (at the present time) to time point N (how much one unit of money invested at time point 0 is worth at time point N) $R_{0\to N,1/N}^{P,N}$ is given by equation (24).

$$R_{0\to N,1/N}^{P,N} = \frac{S_{N,1/N}^{P,N}}{S_{0,1/N}^{P}} = \sum_{i=1}^{4} w_i \prod_{j=1}^{N}\left(1 + X_{j,1/N}^{i}\right) \tag{24}$$

At time point N, the probability that the amount of the portfolio $R_{0\to N,1/N}^{P,N}$ is below the threshold K is given by $P\left(R_{0\to N}^{P,N} \le K\right) = P\left(\sum_{i=1}^{4} w_i \prod_{j=1}^{N}\left(1 + X_{j,1/N}^{i}\right) \le K\right)$. Unlike the case of rebalancing, due to the fact that the weight w_i is outside the stochastic variable $\prod_{j=1}^{N}\left(1 + X_{j,1/N}^{i}\right)$, even though taking logarithm of the portfolio gross return, the downside probability cannot be explicitly derived from the normal distribution.

3.3 Implication when rebalancing frequency equals data sampling frequency

For the case where the rebalancing frequency equals the data sampling frequency, we will again discuss the caution of using a one-period optimization model for the purpose of the long-term portfolio management with rebalancing.

(The case where the frequency is annual)

In the case of annual frequency, putting $\Delta t = 1$ in equation (16), equation (16, $\Delta t = 1$) is attained.

$$X_{k,1}^{i} = \frac{S_{k+1,1}^{i} - S_{k,1}^{i}}{S_{k,1}^{i}} = \mu_1^i \cdot 1 + \sigma_1^i \varepsilon_{k,1}^{i} \tag{16, $\Delta t = 1$}$$

The expected value and standard deviation of the annual return $X_{k,1}^{i}$ are μ_1^i and σ_1^i, respectively. The stochastic variable representing the portfolio gross return from time point 0 (at the present time) to time point T (how much one unit of money invested at time zero is worth at time N) $R_{0\to T,1}^{P}$ is given by equation (25).

$$R_{0\to T,1}^{P} = \frac{S_{T,1}^{P}}{S_{0,1}^{P}} = \prod_{j=1}^{T}\left(1 + \sum_{i=1}^{4} w_i X_{j,1}^{i}\right) \tag{25}$$

At time point T, the probability that the amount of the portfolio $R^P_{0\to N,1/N}$ is below the threshold value K is given by $P\left(R^P_{0\to T,1} \le K\right) = P\left(\prod_{j=1}^{T}\left(1+\sum_{i=1}^{4} w_i X^i_{j,1}\right) \le K\right) = P\left(log\left(\prod_{j=1}^{T}\left(1+\sum_{i=1}^{4} w_i X^i_{j,1}\right)\right) \le logK\right) = P\left(\sum_{j=1}^{T} log\left(1+\sum_{i=1}^{4} w_i X^i_{j,1}\right) \le logK\right) \cong P\left(\sum_{j=1}^{T}\sum_{i=1}^{4} w_i X^i_{j,1} \le logK\right)$. Since the sampling frequency $N=1$ is low because $\sum_{i=1}^{4} w_i X^i_{j,1/N}$ is not sufficiently small, approximation $log\left(1+\sum_{i=1}^{4} w_i X^i_{j,1/N}\right) \cong \sum_{i=1}^{4} w_i X^i_{j,1/N}$ is considered to be rather coarse. In addition, the distribution followed by the stochastic variable $\sum_{i=1}^{4} w_i X^i_{j,1}$ is the normal distribution with the expected value $\boldsymbol{w\mu_1^T}$ and the variance $\boldsymbol{w\Omega_1 w}^T$ composed of the expected values vector and variance-covariance matrices estimated by using T annual return data $\boldsymbol{\mu_1} = \left(\mu_{1,1}, \mu_{2,1}, \mu_{3,1}, \mu_{4,1}\right)$ and $\boldsymbol{\Omega_1}$, respectively. As the numerical example in Section 4 confirms, the confidence intervals for the estimated parameter obtained from the annual return data are considerably wide and the estimation accuracy is quite coarse.

(Cases where frequencies are semi-annual ($\boldsymbol{N=2}$), Quarterly ($\boldsymbol{N=4}$), Monthly ($\boldsymbol{N=12}$), Weekly ($\boldsymbol{N=52}$), daily ($\boldsymbol{N=252}$))

In the case of N times a year frequency, putting $\Delta t = 1/N$ in equation (16), equation (16, $\Delta t = 1/N$) is attained.

$$X^i_{j,1/N} = \frac{S^i_{j+1,1/N} - S^i_{j,1/N}}{S^i_{j,1/N}} = \mu^i_{1/N} \cdot \frac{1}{N} + \sigma^i_{1/N}\varepsilon^i_{j,1/N} \qquad (16, \Delta t = 1/N)$$

The expected values and standard deviations of $1/N$ year return $X^i_{j,1/N}$ are $\mu^i_{1/N}/N$ and $\sigma^i_{1/N}/\sqrt{N}$, respectively. The stochastic variable representing the portfolio gross return from time point 0 (at the present time) to time point $N \cdot T$ (how much one unit of money invested at time point 0 is worth at time point $N \cdot T$) $R^P_{0\to N\cdot T,1/N}$ is given by equation (26).

$$R^P_{0\to N\cdot T,1/N} = \frac{S^P_{N\cdot T,1/N}}{S^P_{0,1/N}} = \prod_{j=1}^{N\cdot T}\left(1+\sum_{i=1}^{4} w_i X^i_{j,1/N}\right) \qquad (26)$$

At time point $N \cdot T$, the probability that the amount of the portfolio $R^P_{0\to N\cdot T,1/N}$ is below the threshold K is given by $P\left(R^P_{0\to N\cdot T,1/N} \le K\right) = P\left(\prod_{j=1}^{N\cdot T}\left(1+\sum_{i=1}^{4} w_i X^i_{j,1/N}\right) \le K\right) = P\left(log\left(\prod_{j=1}^{N\cdot T}\left(1+\sum_{i=1}^{4} w_i X^i_{j,1/N}\right)\right) \le logK\right) = P\left(\sum_{j=1}^{N\cdot T} log\left(1+\sum_{i=1}^{4} w_i X^i_{j,1/N}\right) \le logK\right) \cong P\left(\sum_{j=1}^{N\cdot T}\sum_{i=1}^{4} w_i X^i_{j,1/N} \le logK\right)$. As the sampling frequency N increases, $\sum_{i=1}^{4} w_i X^i_{j,1/N}$ becomes smaller and the accuracy of the approximation $log\left(1+\sum_{i=1}^{4} w_i X^i_{j,1/N}\right) \cong \sum_{i=1}^{4} w_i X^i_{j,1/N}$ is improved. The distribution followed by the stochastic variable $\sum_{i=1}^{4} w_i X^i_{1,1/N}$ is the normal distribution with the expected value $\boldsymbol{w\mu^T}_{\boldsymbol{1/N}}$ and the variance $\boldsymbol{w\Omega_{1/N}w}^T$ with expected values vector and variance-covariance matrices estimated by using $1\,/\,N$-year return data for $N \cdot T$ periods $\boldsymbol{\mu_{1/N}} = \left(\mu_{1,1/N}, \mu_{2,1/N}, \mu_{3,1/N}, \mu_{4,1/N}\right)$ and $\boldsymbol{\Omega_{1/N}}$, respectively.

Thus, when the rebalancing frequency and the data sampling frequency are equal, whether annual, semiannual, monthly, weekly, or daily, the downside probability at a future time can be approximated by a normal distribution. While the confidence intervals for the estimated parameter (the expected value $\boldsymbol{w\mu^T}_{\boldsymbol{1}}$ and the variance $\boldsymbol{w\Omega_1 w}^T$) obtained from the annual return data are considerably large as we mentioned earlier, the confidence intervals for the estimated parameter (the expected value $\boldsymbol{w\mu^T}_{\boldsymbol{1/N}}$ and the variance $\boldsymbol{w\Omega_{1/N}w}^T$) obtained from $1/N$-year return data for $N \cdot T$ periods shrink as the sampling frequency N is increased.

Based on the previous discussion, "the condition that both the frequency of data sampling and the frequency of rebalancing are sufficiently high" should be satisfied to explicitly ascertain the relation between the parameters (mean and variance) of the one-period optimization model and the downside probability from the normal distribution.

4. Numerical examples

In Section 1, we introduced Honda (2019) as a preceding study and pointed out that when attempting optimization based on estimated parameters, the parameter values, in addition to being unbiased, should have their confidence intervals as small as possible, so that a sampling frequency that can ensure a sufficient number of return data is necessary. Also, in Section 3, when the rebalancing frequency and the data sampling frequency are equal, the downside probabilities after T years are approximately given by $P\left(\sum_{j=1}^{T}\sum_{i=1}^{4} w_i X_{j,1}^{i} \leq logK\right)$ (case of annual sampling) and $P\left(\sum_{j=1}^{N\cdot T}\sum_{i=1}^{4} w_i X_{j,1/N}^{i} \leq logK\right)$ (case of $1/N$ year sampling). The confidence interval for the downside probability is affected by a first-order approximation of the logarithmic function as well as the estimated intervals for the expected value and variance of the stochastic variable $\sum_{j=1}^{T}\sum_{i=1}^{4} w_i X_{j,1}^{i}$ and $\sum_{j=1}^{N\cdot T}\sum_{i=1}^{4} w_i X_{j,1/N}^{i}$.

In the numerical example, as return data following a normal distribution, we adopt $1/N$-year sampling return data $X_{j,1/N}$ over $N \cdot T$ period. The sampling frequency can be annual, semiannual, quarterly, monthly, weekly, or daily ($N = 1, 2, 4, 12, 52, 252$), and the years of return data are 5 years and 25 years ($T = 5, 25$), and the confidence intervals for the expected value and standard deviation obtained from these return data sets are examined. For details on the confidence intervals of expected values and standard deviations, please refer to Noda and Miyaoka (1990).

(Confidence interval of expected value)

The expected value $\bar{X}_{1/N}$ and unbiased variance $U_{1/N}^2$ estimated from $1/N$-year sampling return data $X_{j,1/N}$ over $N \cdot T$ periods are given by equations (27) and (28), respectively.

$$\bar{X}_{1/N} = \frac{1}{N \cdot T}\sum_{j=1}^{N\cdot T} X_{j,1/N} \tag{27}$$

$$U_{1/N}^2 = \frac{1}{N \cdot T - 1}\sum_{j=1}^{N\cdot T}\left(X_{j,1/N} - \bar{X}_{1/N}\right)^2 \tag{28}$$

Assumed that the true expected value of $1/N$-year return is $\mu_{1/N}$, $\frac{\bar{X}_{1/N}-\mu_{1/N}}{U_{1/N}^2/\sqrt{N\cdot T}} \sim t_{N\cdot T-1}$, where $t_{N\cdot T-1}$ follows t distribution with degree of freedom $N \cdot T - 1$. If we adopt the upper $100(\alpha/2)$ percentile points, we obtain equation (29).

$$P\left(-t_{N\cdot T-1,\frac{\alpha}{2}} \leq \frac{\bar{X}_{1/N} - \mu_{1/N}}{U_{1/N}^2/\sqrt{N \cdot T}} \leq t_{N\cdot T-1,\frac{\alpha}{2}}\right) = 1 - \alpha \tag{29}$$

Rewriting equation (29), we obtain equation (30).

$$P\left(\bar{X}_{1/N} - t_{N\cdot T-1,\frac{\alpha}{2}}\frac{U_{1/N}}{\sqrt{N\cdot T}} \le \mu_{1/N} \le \bar{X}_{1/N} + t_{N\cdot T-1,\frac{\alpha}{2}}\frac{U_{1/N}}{\sqrt{N\cdot T}}\right) = 1-\alpha \qquad (30)$$

From equation (30), the 100(1 − α) percentile confidence interval of $\mu_{1/N}$ is $\left[\bar{X}_{1/N} - t_{N\cdot T-1,\frac{\alpha}{2}}\frac{U_{1/N}}{\sqrt{N\cdot T}}, \bar{X}_{1/N} + t_{N\cdot T-1,\frac{\alpha}{2}}\frac{U_{1/N}}{\sqrt{N\cdot T}}\right]$. Inequality in the left-hand side of equation (30) multiplied by N yields equation (31).

$$P\left(N\cdot\bar{X}_{1/N} - t_{N\cdot T-1,\frac{\alpha}{2}}\frac{U_{1/N}\sqrt{N}}{\sqrt{T}} \le N\cdot\mu_{1/N} \le N\cdot\bar{X}_{1/N} + t_{N\cdot T-1,\frac{\alpha}{2}}\frac{U_{1/N}\sqrt{N}}{\sqrt{T}}\right) = 1-\alpha \qquad (31)$$

Based on equation (31), the confidence interval for the expected annualized value of $N\cdot\mu_{1/N}$ is given by equation (32).

$$P\left[N\cdot\bar{X}_{1/N} - t_{N\cdot T-1,\frac{\alpha}{2}}\frac{U_{1/N}\sqrt{N}}{\sqrt{T}}, N\cdot\bar{X}_{1/N} + t_{N\cdot T-1,\frac{\alpha}{2}}\frac{U_{1/N}\sqrt{N}}{\sqrt{T}}\right] \qquad (32)$$

Noticing that the equation that $N\cdot\bar{X}_{1/N}$ satisfies is $1\cdot\bar{X}_1 = 2\cdot\bar{X}_{1/2} = 4\cdot\bar{X}_{1/4} = 12\cdot\bar{X}_{1/12} = 52\cdot\bar{X}_{1/52} = 252\cdot\bar{X}_{1/252}$ and also the equation that $U_{1/N}\sqrt{N}$ satisfies is $U_1\sqrt{1} = U_{1/2}\sqrt{2} = U_{1/4}\sqrt{4} = U_{1/12}\sqrt{12} = U_{1/52}\sqrt{52} = U_{1/252}\sqrt{252}$, the essential influence of the difference in sampling frequency on the confidence interval of the annualized expected value is the degree of freedom of the confidence coefficient $t_{N\cdot T-1,\frac{\alpha}{2}}$.

(Confidence interval of variance)

Assumed that the variance of $1/N$-year return is $\sigma^2_{1/N}$, $\frac{\sum_{j=1}^{N\cdot T}(X_{j,1/N}-\bar{X}_{1/N})^2}{\sigma^2_{1/N}} \sim \chi^2_{N\cdot T-1}$, where $\chi^2_{N\cdot T-1}$ follows χ^2 distribution with degree of freedom $N\cdot T-1$. If we adopt the upper 100($\alpha/2$) percentile points, we obtain equation (33).

$$P\left(\chi^2_{N\cdot T-1,1-\alpha/2} \le \frac{\sum_{j=1}^{N\cdot T}\left(X_{j,1/N}-\bar{X}_{1/N}\right)^2}{\sigma^2_{1/N}} \le \chi^2_{N\cdot T-1,\alpha/2}\right) = 1-\alpha \qquad (33)$$

Rewriting equation (33), we obtain equation (34).

$$P\left(\frac{\sum_{j=1}^{N\cdot T}\left(X_{j,1/N}-\bar{X}_{1/N}\right)^2}{\chi^2_{N\cdot T-1,\alpha/2}} \le \sigma^2_{1/N} \le \frac{\sum_{j=1}^{N\cdot T}\left(X_{j,1/N}-\bar{X}_{1/N}\right)^2}{\chi^2_{N\cdot T-1,1-\alpha/2}}\right) = 1-\alpha \qquad (34)$$

From equation (34), the confidence interval for $\sigma^2_{1/N}$ is $\left[\frac{\sum_{j=1}^{N\cdot T}(X_{j,1/N}-\bar{X}_{1/N})^2}{\chi^2_{N\cdot T-1,\alpha/2}}, \frac{\sum_{j=1}^{N\cdot T}(X_{j,1/N}-\bar{X}_{1/N})^2}{\chi^2_{N\cdot T-1,1-\alpha/2}}\right]$.

After the left-hand side of equation (34) is rewritten by way of equation (28), the resulting inequality is multiplied by N, then, the confidence interval of the annualized variance $N\cdot\sigma^2_{1/N}$ is obtained as equation (35).

$$P\left(\frac{(N\cdot T-1)}{\chi^2_{N\cdot T-1,\alpha/2}}N\cdot U^2_{1/N} \le N\cdot\sigma^2_{1/N} \le \frac{(N\cdot T-1)}{\chi^2_{N\cdot T-1,1-\alpha/2}}N\cdot U^2_{1/N}\right) = 1-\alpha \qquad (35)$$

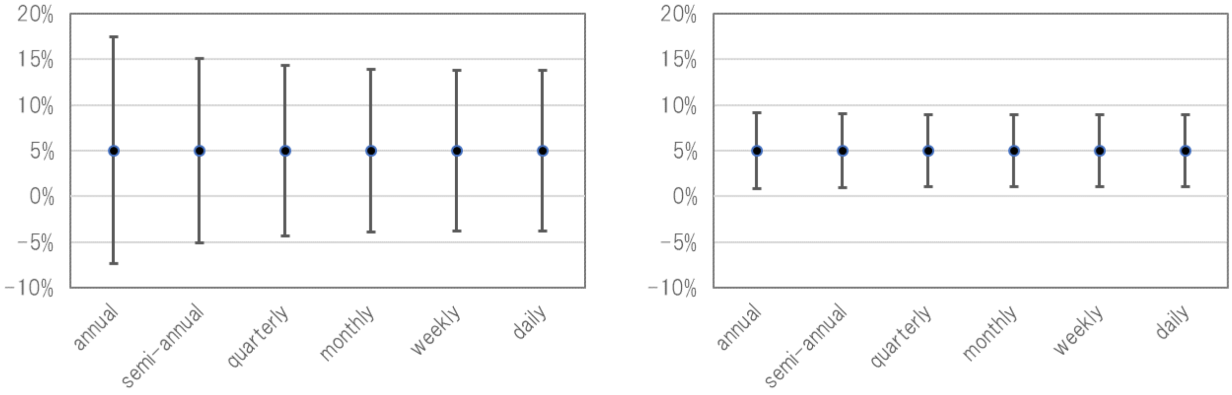
Noticing that the equation that $N\cdot U^2_{1/N}$ satisfies is $1\cdot U_1^2 = 2\cdot U^2_{1/2} = 4\cdot U^2_{1/4} = 12\cdot U^2_{1/12} = 52\cdot U^2_{1/52} = 252\cdot U^2_{1/252}$, the essential influence of the difference in sampling frequency on the confidence interval of the annualized variance is values of the coefficients $\frac{(N\cdot T-1)}{\chi^2_{N\cdot T-1,\alpha/2}}$ and $\frac{(N\cdot T-1)}{\chi^2_{N\cdot T-1,1-\alpha/2}}$.

True expected value $N\cdot\mu_{1/N}$ and standard deviation $\sqrt{N}\cdot\sigma_{1/N}$ of the annual return are set to 5% and 10%, respectively, and also confidence level α is set to 0.05. The confidence intervals for the expected value and standard deviation are shown in Figure 1 and Figure 2, respectively, for 5 and 25 years of return data. The confidence intervals are shown for the case where the estimated expected value and the variance are equal to the true expected value and the variance of the return ($\mu_{1/N} = \bar{X}_{1/N}, \sigma_{1/N}^2 = U_{1/N}^2$).

Focusing on the confidence interval of the expected value, we find that the confidence interval shrinks significantly when the number of years T of return data is increased from 5 years (Figure 1a) to 25 years (Figure 1b). However, the effect of increasing the sampling frequency on the confidence interval is small, especially when the number of years of return data is 25 years.

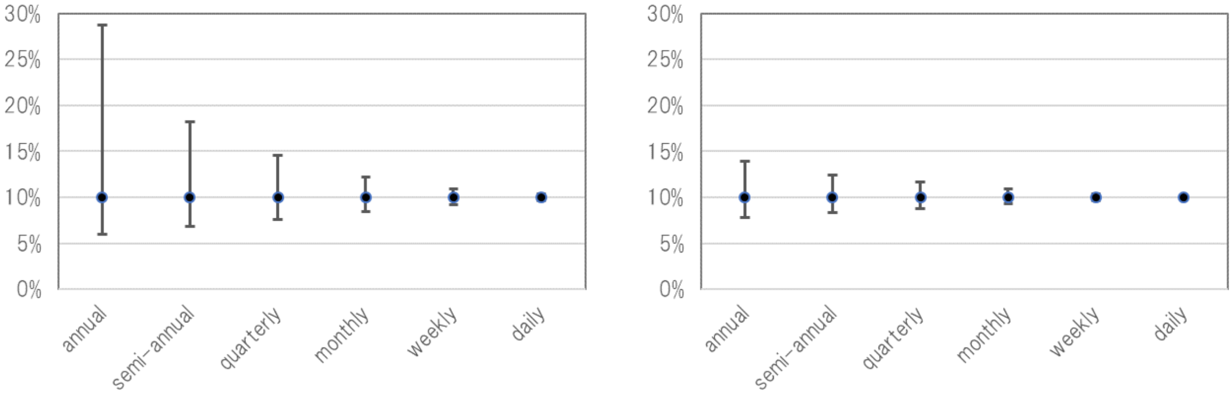
Regarding the confidence interval of the standard deviation, if the number of years of return data T is increased from 5 years (Figure 2a) to 25 years (Figure 2b), the confidence interval shrinks significantly, as in the case of expected values. As the sampling frequency N is increased, the confidence interval shrinks regardless of the number of years of return data. This effect is particularly large when the number of years of return data is 5 years.

From the numerical example, two implications for the case where investment opportunities do not vary are as follows. First, the confidence interval of the standard deviation shrinks rapidly as the sampling frequency increases, so to grasp the downside probability sharply, it is effective to increase the data sampling frequency within a realistic rebalancing frequency. Next, the confidence interval for the expected value will not shrink essentially unless a longer number of years of return data is taken. However, there is concern that attempting to take a longer number of years of return data may include periods of time that may differ from future economic assumptions. It would be preferable to use an approach other than simply estimating expected returns from historical return data, such as using forward-looking expected returns.



(a) Years of return data T=5 (b) Years of return data T=25

Fig.1 95% confidence intervals for annualized expected values



(a) Years of return data T=5 (b) Years of return data T=25

Fig. 2 95% confidence intervals for annualized standard deviations

5. Summary and future issues

In this paper, we show that for both continuous and discrete models, when investment opportunities do not vary, "both the frequency of data sampling and the frequency of rebalancing are sufficiently high" should be satisfied to explicitly capture the relation between parameters (mean and variance) of one-period optimization model and the downside probability from the normal distribution. In addition, as an implication from the viewpoint of capturing the efficient frontier and downside probability sharply, it was found that it is better to collect return data at a sampling frequency that can take as many samples as possible.

Here are four main issues to be addressed in the future.

i. The target amount for determining the downside probability was set as a constant value K. However, it is desirable to extend the model to include the target amount as a stochastic variable in the application of the model.
ii. Explore appropriate estimation methods, market data, etc. regarding forward-looking expected returns and variance-covariance matrices.
iii. Based on the empirical analysis, extend the model to the case of variable investment opportunities, if necessary.
iv. The GPIF's policy asset mix is reviewed every five years based on the fiscal verification. The appropriateness of using a one-period model for long-term investment would need to be fully considered. In this case, the framework of the multi-period portfolio optimization model in Hibiki (2004) may be useful.

References

[1] Shimizu, Y. and T. Uchiyama (2017), "Is There a Rebalancing Premium?" Working Paper, https://researchmap.jp/tomonori.uchiyama/misc/43929355/attachment_file.pdf (In Japanese)

[2] Hibiki, N. (2004), "A Multi-Period Portfolio Optimization Model for Pension ALM." Mizuho Pension Report, May/June issue, pp. 36-48. https://lab.ae.keio.ac.jp/~hibiki_lab/profile_2/MPR_n55.pdf (In Japanese)

[3] Honda, T. (2019), "Portfolio Selection Theory in its 65th Year." Contemporary Finance, 40, pp. 1-24.

[4] Government Pension Investment Fund (GPIF) (2023), Annual Report. (In Japanese)

[5] Noda, K. and E. Miyaoka, (1990), "Introduction and Exercises in Mathematical Statistics." Kyoritsu Shuppan. (In Japanese)

[6] Miyazaki, K. (2005), "Box-Cox transformation and option evaluation." Applied Statistics, 34(2), pp. 75-97. (In Japanese)

[7] Morimura, H. and M. Kijima (1991), "Stochastic Processes for Finance." JUSE. (In Japanese)

[8] Kan, R. and Zhou, G. (2007), "Optimal Portfolio Choice with Parameter Uncertainty." Journal of Financial and Quantitative Analysis, 42(3), 621-656.

[9] Levy, M. and Roll, R. (2010) "The market portfolio may be mean/variance efficient after all." Review of Financial Studies, 23 (6), 2464-2491.

[10] Markowitz, H. (1952) "Portfolio Selection." The Journal of Finance, 7(1), 77-91.