DYNAMIC ASSET PORTFOLIO OPTIMIZATION BY A HEURISTIC GA-BASED METHOD

Y. ORITO†

Department of Systems and Information Engineering, Ashikaga Institute of Technology
268-1, Ohmae-cho, Ashikaga, Tochigi 326-8558, Japan

Y. KAMBAYASHI and Y. TSUJIMURA

Department of Computer and Information Engineering, Nippon Institute of Technology
4-1, Gakuendai, Miyashiro-cho, Saitama 345-8501, Japan

M. INOGUCHI and H. YAMAMOTO

Faculty of System Design, Tokyo Metropolitan University
6-6, Asahigaoka, Hino, Tokyo 191-0065, Japan

The portfolio optimizations are generally to optimize the proportion of funds in the static asset portfolio. Then, we have a problem that the assets initially not included in the portfolio are never optimized. In order to avoid this problem, we propose a GA-based method that optimizes the portfolio that consists of not only the given static assets but also the dynamic assets in this paper. In order to demonstrate the effectiveness of our method, we apply this method to the Tokyo Stock Exchange to create an index fund. This fund has passive portfolios. The numerical experiments show that our method works well for a dynamic asset portfolio optimization.

1. Introduction

The portfolio optimization problem, based on the Markowitz's modern portfolio theory [1], is one of the combinatorial optimization problems that select assets for a certain portfolio in advance and to optimize the proportion of funds in these static assets. There are many reports using evolutional algorithms to solve this problem (for this, see e.g., [2-4]). However, the static assets included in the portfolios have already been determined before applying optimization methods. Therefore we have a critical problem that the assets not included in the portfolio at the first place are never optimized. On the other hand, Orito et al. [5] proposed a method to construct a portfolio by selecting a subset of the given

† Work partially supported by Grant #20710119, Grant-in-Aid for Young Scientists (B) from Japan Society for the Promotion of Science, (2008-).

assets. Aranha and Iba [6] also proposed a similar but different method. Both methods employ the binary digit 0 or 1 as the representations of "not-selected" or "selected" assets. However, when the number of selected assets for the portfolios increases, the number of combinations of the proportion of funds also increases. Therefore it is hard to optimize the proportion of funds in the portfolio consisting of the static large number of assets. In order to avoid this problem, we propose Heuristic GA-based Method that makes portfolios consisting of not the given static assets but the dynamically selected assets in this paper.

2. Portfolio Optimization Problem

We define the following notations for the dynamically selected asset portfolio optimization.

i : Asset i , $i = 1, \cdots, N$.

$\mathbf{X}$: the set consisting of assets. That is $\mathbf{X} = \{i_1, \cdots, i_{\#\mathbf{X}}\}$ whose $\#\mathbf{X}$ represents the number of elements in the set $\mathbf{X}$.

$f_{index}(t)$: the value of the benchmark index at t .

$\mathbf{P}_{index}$: the the rates of changes of benchmark index over $t = 1, \cdots, T$. That is the vector $\mathbf{P}_{index} = (P_{index}(1), \cdots, P_{index}(T))$ whose $P_{index}(t)$ is defined by $(f_{index}(t+1) - f_{index}(t)) / f_{index}(t)$.

g_i : the weight of Asset i included in a portfolio. That is $0 \le g_i \le 1$.

$\mathbf{G}_\mathbf{X}$: the portfolio for the set $\mathbf{X}$. That is vector $\mathbf{G}_\mathbf{X} = (g_{i_1}, \cdots, g_{i_{\#\mathbf{X}}})$ such that $\sum g_{i_j} = 1$.

$f_{\mathbf{G}_k}(t)$: the value of the portfolio $\mathbf{G}_\mathbf{X}$ at t . This value means the portfolio price multiplied by the number of stocks.

$\mathbf{P}_{\mathbf{G}_\mathbf{X}}$: the return rates of portfolio $\mathbf{G}_\mathbf{X}$ over $t = 1, \cdots, T$. That is the vector $\mathbf{P}_{\mathbf{G}_\mathbf{X}} = (P_{\mathbf{G}_\mathbf{X}}(1), \cdots, P_{\mathbf{G}_\mathbf{X}}(T))$ where $P_{\mathbf{G}_\mathbf{X}}(t)$ is defined by $(f_{\mathbf{G}_\mathbf{X}}(t+1) - f_{\mathbf{G}_\mathbf{X}}(t)) / f_{\mathbf{G}_\mathbf{X}}(t)$.

In this paper, we apply the optimization method to index funds. Index funds are popular passive portfolios and are constructed to mimic the performance of the given benchmark index, such as the S&P 500 in New York. As an objective function to evaluate the relation between the fund's price and the benchmark index, we adopt the correlation coefficient between the fund's return rates and the rates of changes of the benchmark index as our objective function,

$$\max \quad R_{\mathbf{G}_\mathbf{X}} = \frac{\mathrm{cov}(\mathbf{P}_{index}, \mathbf{P}_{\mathbf{G}_\mathbf{X}})}{\sqrt{\mathrm{var}(\mathbf{P}_{index}) \cdot \mathrm{var}(\mathbf{P}_{\mathbf{G}_\mathbf{X}})}} . \tag{1}$$

3. Heuristic GA-based Method

Suppose that the initial set consisting of N assets is defined as a set $\mathbf{A} = \{1, \cdots, N\}$. Based on this set $\mathbf{A}$, two subsets are defined as $\mathbf{H}(\subseteq \mathbf{A})$ and $\mathbf{L}(\subseteq \mathbf{A})$. We note that the subset $\mathbf{H}$ is a subset consisting of assets selected for the portfolio and the subset $\mathbf{L}$ is a subset consisting of assets not selected for the portfolio. Hence, as the relation between the subsets $\mathbf{H}$ and $\mathbf{L}$, note that $\mathbf{A} = \mathbf{H} \bigcup \mathbf{L}$ and $\mathbf{H} \bigcap \mathbf{L} = \phi$ hold.

We first initialize $\mathbf{L} = \mathbf{A}$, and then the Heuristic GA-based Method alternately repeats the following two steps: Steps A and B. Step A is to move the assets expected to be greatly good influence on the objective function from the subset $\mathbf{L}$ into the subset $\mathbf{H}$. Step B is to remove the assets expected to be no-good influence on the objective function from the subset $\mathbf{H}$ into the subset $\mathbf{L}$. The GA in our method is described in Section 3.1. Steps A and B are described in Section 3.2. The procedure of our method is shown in Fig. 1.

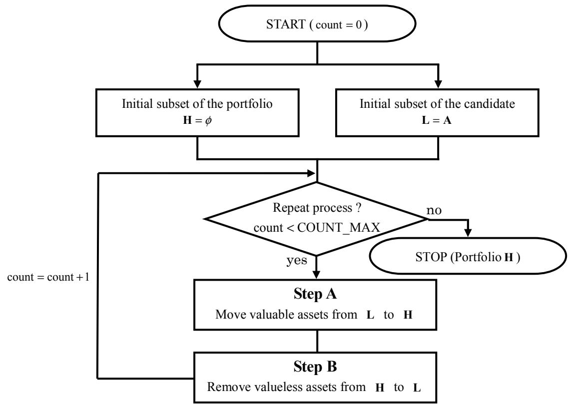

Figure 1. The procedure of Heuristic GA-based Method consisting of Steps A and B.

3.1. *GA*

In Heuristic GA-based Method, we apply the following GA to the subsets $\mathbf{H}$ and $\mathbf{L}$ in parallel for Step A and to the subset $\mathbf{H}$ for Step B, respectively.

For the genetic representation, a gene represents the weight of Asset i, and is denoted by $g_i\ (0 \le g_i \le 1)$. A chromosome represents a portfolio for the set $\mathbf{X}$, and is denoted by $\mathbf{G_X} = (g_{i_1}, \cdots, g_{i_{\#\mathbf{X}}})$. The "*fitness value of GA*" is to maximize the correlation coefficient given by $R_{\mathbf{G_X}}$.

For the genetic operations, the GA randomly generates M chromosomes for the initial population. In the above genetic representation, the chromosome

$\mathbf{G_X} = \left(g_{i_1}, \cdots, g_{i_{\#\mathbf{X}}}\right)$ consists of the dynamically selected genes. The order of genes in the chromosome is not significant. Hence, we apply the uniform crossover and the uniform mutation to make a new offspring. After making a new offspring, the GA repairs the new genes via renormalization. On each generation, the GA selects M chromosomes by using the elitism selection in the order of high fitness values of GA. The GA repeats these operations until the terminate criterion, the final K -th generation, is satisfied. On the K -th generation, we choose one chromosome with the highest fitness value of GA in the population. This chromosome is expressed as $\mathbf{G_X}*$ for the subset $\mathbf{X}$.

3.2. *Steps A and B in Heuristic GA-based Method*

Step A in our method is to move the assets expected to be greatly good influence on the correlation coefficient from the subset $\mathbf{L}$ into the subset $\mathbf{H}$.

Let s be the number of repetition of the procedure for Step A. We define $\mathbf{H}_s$ and $\mathbf{L}_s$ as subsets $\mathbf{H}$ and $\mathbf{L}$ on the s -th repetition of the procedure, respectively. First, the subsets $\mathbf{L}_0$ and $\mathbf{H}_0$ are initialized as $\mathbf{L}_0 = \mathbf{A}$ and $\mathbf{H}_0 = \phi$. On the K -th generation of GA for the s -th procedure, we define the subset $\mathbf{J}_{s+1}$ as the group of assets that moves from $\mathbf{L}_s$ to $\mathbf{H}_{s+1}$. Our method moves the assets whose weight belongs to the chromosome $\mathbf{G}_{\mathbf{L}_s}*$ from the subset $\mathbf{L}_s$ to the subset $\mathbf{H}_{s+1}$ by using the heuristic rule defined as

$$\begin{aligned}
\mathbf{J}_{s+1} &= \{i \mid g_i \geq B_A, i \in \mathbf{L}_s\} \\
\mathbf{H}_{s+1} &= \mathbf{H}_s \cup \mathbf{J}_{s+1} = \bigcup_{j=0}^{s+1} \mathbf{J}_j \ , \\
\mathbf{L}_{s+1} &= \mathbf{L}_s \setminus \mathbf{J}_{s+1}
\end{aligned} \tag{2}$$

where the boundary parameter B_A is given in advance.

We have the following property that the coefficient $R_{\mathbf{G}_{\mathbf{H}_s}}$, which is obtained by the portfolio $\mathbf{G}_{\mathbf{H}_s}$ consisting of the union sets $\mathbf{H}_s = \mathbf{J}_0 \cup \cdots \cup \mathbf{J}_s$, is greater than or equal to the minimum of the sequence of coefficients, $R_{\mathbf{G}_{\mathbf{J}_0}}$ to $R_{\mathbf{G}_{\mathbf{J}_s}}$.

Property.

If $f_{\mathbf{G}_{\mathbf{X}_a}}(t) \big/ \left(f_{\mathbf{G}_{\mathbf{X}_a}}(t) + f_{\mathbf{G}_{\mathbf{X}_b}}(t)\right)$ *is a constant for* $\forall t \geq 0$ *and* $R_{\mathbf{G}_{\mathbf{X}_a}} \geq 0$ *and* $R_{\mathbf{G}_{\mathbf{X}_b}} \geq 0$ *are satisfied, the relation of correlation coefficients obtained by the portfolios is defined as follows.*

$$R_{\mathbf{G}_{\mathbf{X}_a \cup \mathbf{X}_b}} \geq \min\left(R_{\mathbf{G}_{\mathbf{X}_a}}, R_{\mathbf{G}_{\mathbf{X}_b}}\right).$$

From the property, we can define the subset $\mathbf{H}_{s+1}$ as the union set of $\mathbf{H}_s$ and $\mathbf{J}_{s+1}$. However, we cannot obtain the accurate number of assets that should move to the portfolio only by the theorem. Hence, we define Eq. (3) as the terminate criterion of procedure of Step A.

$$R_{\mathbf{G}_{\mathbf{H}_s}*} \leq R_{\mathbf{G}_{\mathbf{H}_{s-1}}*}. \tag{3}$$

When the terminate criterion is satisfied, we get the chromosome $\mathbf{G}_{\mathbf{H}_{s-1}}*$ for the subset $\mathbf{H}_{s-1}$ as our portfolio obtained by Step A.

On the other hand, Step B in our method is to remove the assets expected to be no-good influence on the correlation coefficient from the subset $\mathbf{H}$ into $\mathbf{L}$.

Let q be the number of repetition of the procedure for Step B. We define $\mathbf{H}_q$ and $\mathbf{L}_q$ as subsets $\mathbf{H}$ and $\mathbf{L}$ on the q-th repetition of the procedure, respectively. First, the subsets $\mathbf{L}_0$ and $\mathbf{H}_0$ are initialized as the subsets obtained by Step A, $\mathbf{L}_0 = \mathbf{L}_{s-1}$ and $\mathbf{H}_0 = \mathbf{H}_{s-1}$. On the K-th generation of GA for the q-th procedure, we define the subset $\mathbf{J}_{q+1}$ as the group of assets that moves from $\mathbf{H}_q$ to $\mathbf{L}_{q+1}$. Our method removes the assets whose weight belongs to the chromosome $\mathbf{G}_{\mathbf{H}_q}*$ from the subset $\mathbf{H}_q$ to the subset $\mathbf{L}_{q+1}$ by using the heuristic rule defined as

$$\begin{aligned} &\mathbf{J}_{q+1} = \{i \mid g_i \leq B_B, i \in \mathbf{L}_q\} \\ &\mathbf{H}_{q+1} = \mathbf{H}_q \setminus \mathbf{J}_{q+1} \\ &\mathbf{L}_{q+1} = \mathbf{L}_q \bigcup \mathbf{J}_{q+1} \end{aligned}, \tag{4}$$

where the boundary parameter B_B is given in advance.

As the terminate criterion of the procedure of Step B, we define Eq. (5).

$$R_{\mathbf{G}_{\mathbf{H}_q}*} \geq R_{\mathbf{G}_{\mathbf{H}_0}*}. \tag{5}$$

When the terminate criterion is satisfied, we get the chromosome $\mathbf{G}_{\mathbf{H}_q}*$ for the subset $\mathbf{H}_q$ as our portfolio obtained by Step B.

As shown in Fig. 1, our Heuristic GA-based Method repeats Steps A and B $\mathrm{COUNT\_MAX}$ times. Finally, we can get the dynamic asset portfolio $\mathbf{G}_{\mathbf{H}_q}*$.

4. Numerical Experiments

We have applied our method to each of 12 data periods of the First Section of Tokyo Stock Exchange from Jan. 6, 1997 to Oct. 2, 2006. Each data period is 100 days, and is shifted every 200 days. The dataset is a subset of the TOPIX (Tokyo Stock Price Index). The TOPIX is a well known benchmark index and represents the increase or decrease in stock values of all assets on the market.

The parameters used in our method are as follows: Total number of assets: N = 1000, Population size M = 100, Crossover rate = 0.9, Mutation rate = 0.1, Generation size K = 100, Method run = 20, Boundary parameter in Eq. (2) B_A = 40 and Boundary parameter in Eq. (4) B_B = 40.

In order to demonstrate the efficiency of our method, we compare three methods as follows: **Method 1** is our proposed method consisting of Steps A and B, **Method 2** is a compared method consisting only of Step A, and **Method 3** is a compared method that applies the GA to optimize the weight of the static asset portfolio consisting of 1000 ($= N$) assets. For each period, the best correlation coefficients obtained by the three methods are shown in Table 1. In order to investigate of the comparison between the two dynamic asset portfolios obtained by Methods 1 and 2, the numbers of assets selected for the portfolios by the two methods are shown in Tables 2.

Table 1. The best correlation coefficient obtained by Method 1, 2 or 3.

Data Period	Method 1	Method 2	Method 3
1	0.999952	0.999880	0.998378
2	0.999966	0.999881	0.998237
3	0.999976	0.999950	0.998890
4	0.999928	0.999781	0.997120
5	0.999474	0.999011	0.989048
6	0.999907	0.999770	0.994800
7	0.999946	0.999698	0.997371
8	0.999958	0.999876	0.998263
9	0.999948	0.999806	0.995843
10	0.999989	0.999950	0.999128
11	0.999985	0.999934	0.998178
12	0.999983	0.999943	0.999441

We can observe that the coefficients obtained by Methods 1 are higher than those of Methods 2 and 3. Furthermore, Table 2 shows that the standard deviations (SD) of the numbers of assets selected to the portfolios obtained by execution of Method 1 for 20 times are smaller than those of Method 2 in all periods except Periods 3, 11 and 12 (But, there is hardly any difference of SDs for these three periods). This means that Step B in our method works well to remove valueless assets from the portfolios obtained by the process of Step A.

Table 2. The number of assets selected for the dynamic asset portfolio obtained by Method 1 or 2.

Data Period	Method 1				Method 2			
	Largest	Smallest	Avg.	SD	Largest	Smallest	Avg.	SD
1	382	265	326.55	26.6	476	233	341.6	52.3
2	325	258	288	19.6	394	276	339	39.9
3	370	249	305.8	32.9	389	291	334.65	29.8
4	390	84	312.35	79.5	476	9	342.95	119.8
5	335	181	239.25	31.3	309	141	222.95	41.4
6	322	254	289	20.8	351	204	278.1	37.2
7	366	255	292.5	27.4	419	221	327.85	48.9
8	379	252	299.55	29.2	395	250	315.2	38.9
9	343	245	285.9	22.6	375	217	291.9	41.5
10	381	285	330.6	34.0	472	238	375.35	59.5
11	393	250	324.55	38.4	444	280	388.65	36.4
12	442	317	375.9	38.2	479	366	426.15	36.8

5. Conclusions

In this paper, we have proposed a Heuristic GA-based Method for the dynamic asset portfolio optimization. The numerical experiments demonstrate that our method works well for the optimization problem that makes the portfolios consisting of the dynamically selected small number of assets. Our next object is to stabilize the number of dynamically selected assets.

References

1. H. Markowitz, Portfolio selection, *Journal of Finance*, **7**, 77-91 (1952).
2. T.J. Chang, N. Meade, J.E. Beasley and Y.M. Sharaiha, Heuristics for cardinality constrained portfolio optimization, *Computers & Operations Research*, **27**, 1271-1302 (2000).
3. C.C. Lin and Y.T. Liu, Genetic algorithms for portfolio selection problems with minimum transaction lots, *European Journal of Operational Research*, **185-1**, 393-404 (2008).
4. F. Streichert and M. Tanaka-Yamazaki, The Effect of Local Search on the Constrained Portfolio Selection Problem, *Proc. of IEEE Congress on Evolutionary Computation*, 2368-2374 (2006).
5. Y. Orito, H. Yamamoto, and G. Yamazaki, Index Fund Selections with Genetic Algorithms and Heuristic Classifications, *Journal of Computers & Industrial Engineering*, **45**, 97-109 (2003).
6. C. Aranha and H. Iba, Portfolio Management by Genetic Algorithms with Error Modeling, *JCIS Online Proc. of Computational Intelligence in Economics & Finance* (2007).