Pak. J. Statist.
2017 Vol. 33(6), 441-448

PORTFOLIO SELECTION USING ANT COLONY ALGORITHM AND ENTROPY OPTIMIZATION

Mohamadtaghi Rahimi[1§], Pranesh Kumar[2] and **Gholamhossein Yari[1]**
[1] Department of Mathematics, Iran University of Science and Technology Tehran, Islamic Republic of Iran. Email: Yari@iust.ac.ir
[2] Department of Mathematics and Statistics University of Northern British Columbia, Prince George, BC, Canada Email: Pranesh.Kumar@unbc.ca
[§] Corresponding author Email: mt_rahimi@alumni.iust.ac.ir

ABSTRACT

The main objective of portfolio management was to select a suitable stock portfolio. In this study, we used a novel model based on ant colony algorithm and entropy optimization to select an optimal portfolio. Meanwhile we introduced a new definition of covariance to consider a difference between the kinds of influences that the shares have on each other. We first reviewed the literature as well as different studies on the subject, and found the four the most important criteria to be more suited for our research objectives. Of all companies listed at Tehran Stock Exchange, data of 55 companies from 2003 to 2013 were analyzed as the statistical sample; these companies had the highest return on assets and the lowest risk level. Ultimately, we predicted several portfolios containing the best companies with respect to the mentioned criteria. The MATLAB software application was then used for data analysis, and after developing the entropy decision matrix, portfolios with the lowest entropy were selected.

KEYWORDS

Portfolio, Ant colony algorithm, Entropy, Covariance+, Decision matrix.

JEL Classification: G11, 54C70

1. INTRODUCTION

Complicated mechanisms and feedbacks of financial markets distinguish them from other complex systems. Every day, many people put a lot of effort to assess and analyze stocks of different financial markets, and this has led to creation of novel methods, which in combination with classical methods, are being used every day to find further solutions for maximizing profit.

Selecting the portfolio is one of the most important fields of financial decision making, where the presence of uncontrollable variables has significantly influenced the decision making process. This is of high importance for final investors for allocating their budgets towards financial assets in the investment portfolio. Identifying the factors involved in an investor's decision making, on the one hand, and measuring these factors on the other as well as their effects on selecting the portfolio are among the main problems of financial analysts.

As mentioned earlier, financial markets are uncertain and since investors are looking for increasing their revenue and reducing their risk, diversification and creating a portfolio can give the investors some certainty. In this section, the investor or the decision maker must select and evaluate stocks available in the stock exchange, which are considered as investment opportunities.

2. RESEARCH LITERATURE

So far, there have been numerous studies on how to select a portfolio, and different patterns have been suggested for solving the problem of portfolio management which are designed based on different conditions and restrictions.

Markowitz (1952) was the first scholar to propose the modern theory of portfolio in 1952, which argues that we can create a set of portfolios so that considering a certain risk, each stock from that set has the highest expected revenue and yield rate.

However, these linear models, including the Markowitz model, are not capable to accurately determine non-linear behavior and they are only able to determine the linear part of the behavior. Therefore, we need non-linear patterns and models for determining stock behavior and these models have a significant impact on predicting the future of stocks and making the correct decisions regarding these stocks.

In 1973, Sang and Leru published a study entitled “Selecting Optimal Portfolio for Investment Firms”. In this study, 61 firms were evaluated using criteria including beta, returns, C, D, and risk. Finally, by ranking these criteria, an optimal portfolio based on the planning model was proposed.

Lazu et al. (2000) used the genetics algorithm and neural networks for selecting the management for a set of assets and the results showed that the returns of the proposed portfolio obtained from the model was identical to the market index and in some cases, it was even higher.

Yan and Miayu (2007) used a combination of PSO and GA methods to select a multi-period portfolio using the half-variance risk factor. They showed that using both methods together is significantly better than using only one of them.

Chiam et al. (2008) used an imitation pattern or a swarm algorithm for the applications of financial computations. Their objective in this study was to predict the time series of stock price and to develop an optimal portfolio. The results of the study proved that stock prices have chaotic fluctuations and using the swarm algorithm, algorithm teaching methods, testing methods and evaluating the swarm algorithm on a structural basis, we could provide a pattern with the lowest possible prediction error and the highest possible accuracy.

Chang and Lee (2012) investigated the issue of selecting a suitable portfolio of projects. They focused on solving the problem of limitations faced by firms when utilizing capital resources. Accordingly, they used a model based on data envelopment analysis (DEA) for solving this problem. Using this model and bee colony algorithm in artificial intelligence, a comparative process on the problem of optimization in uncertain problems of the industry was carried out.

Najafi (2014) published a study entitled "Optimization of Portfolio Using And Colony Algorithm and the Grey Algorithm". In this study, by evaluating 150 firms active in the stock exchange which had the highest Return On Assets (ROA), a model based on ant colony algorithm and grey algorithm was developed. Using this model they predicted the stocks of those types of firms. They concluded that ant colony algorithm, the grey algorithm, and finally the Markowitz model had the highest accuracy, respectively.

Besides, Entropy as a measure of diversity is recently employed in various sciences. For example, Yari et al. (2017) employed it in option pricing and Farnoosh et al. (2016) proposed a method for image processing by intuitionistic fuzzy entropy. There have been also some improvements in the conceptions of entropy in the last decade. Some of them can be found in Markechová et al. (2016) and Yari et al. (2016).

2.1 Reviewing Concepts and Definitions

2.1.1 Ant Colony Optimization

Ant colony algorithm is one of the most successful swarm-based algorithms. This algorithm is a swarm intelligence algorithm based on the searching behavior of ants in nature. The most interesting trait of ant behavior is their ability to find the paths inside the colony and food sources by surveying pheromone tracks. Then, ants use a pheromone-based probabilistic decision to select the next path. The stronger the pheromone track, the higher its utility. Since ants use the remaining pheromone to follow the path, this behavior triggers a self-help process for creating defined paths by the high concentration of pheromone. By modeling and simulating the heuristic behavior of ants, ordering the eggs, building the colony and storing food and so on, algorithms can be developed to be used for solving complex and combinatory optimization problems. The first algorithm called the ant system was developed by Dorigo et al (1996-1997) and was successfully tested for the comparative evaluation of the traveling salesman problem.

Ant colony algorithm is classified as follows in three main functions (2006).

Auto Solutions Construct: this function carries out the process of developing solutions where due to the neighboring states of a problem, the ants must move based on a transfer rule.

Pheromone Update: this function involves updating the pheromone. Besides reinforcing the pheromone, ant colony algorithm involves the evaporation of the pheromone tracks. Evaporation of pheromone tracks helps ants to forget bad solutions previously learned.

Daemon Actions: this is an optional step in the algorithm which involves applying redundant updates from a general point of view. This function may involve additional reinforcement of the pheromone for the best developed solution.

2.1.2 Entropy

Entropy is a concept used in social sciences, physics, and the information theory for measuring the disorder and uncertainty. In the decision matrix, entropy can also show the dispersion of the index values.

If a variable takes the value of X_i with the probability, P_i, the discrete entropy is calculated by the following equation:

$$E = S(P_1, P_2, \dots, P_m) = -k \sum_{i=1}^{m} P_i . Ln(P_i) \quad k = \frac{1}{Ln(m)}$$

Entropy value of each indicator in the decision matrix is obtained using the following formula:

$$E_j = -k . \sum_{i=1}^{m} P_{ij} . Ln\left(P_{ij}\right) \quad P_{ij} = \frac{a_{ij}}{\sum_{i=1}^{m} a_{ij}}$$

$$D = \begin{bmatrix} a_{11} \; a_{i1} \; a_{1n} \\ a_{21} \; a_{2n} \\ \dots \dots \; \dots \; \dots \dots \\ a_{i1} \; a_{ij} \; a_{in} \\ \dots \dots \; \dots \dots \\ a_{m1} \; a_{mn} \end{bmatrix}_{(m \times n)}$$

Having determined the entropy value of each indicator, dispersion of the values in each indicator j is obtained from the following formula:

$$w_j = \frac{d_j}{\sum_{j=1}^{n} d_j} \tag{2.1}$$

$$w_j = \frac{\lambda_j d_j}{\sum_{j=1}^{n} \lambda_j d_j} \tag{2.2}$$

If there is no verdict about the relative weight of the indicators, the weights of the indicators can be obtained from Equation 2.1 and if the weight of each indicator j is determined as λ_j, the weight of each indicator will be obtained from Equation 2.2, where di is the distance between the edges.

3. PORTFOLIO OPTIMIZATION MODEL

Before we discuss the model, we need to introduce a theorem to measure the risk between shares.

Theorem 3.1

In selecting the portfolio, the risk due to selecting firm j after firm i is defined as follows:

$$Risk(j|i) = \frac{Cov(i,j)^{+}}{Cov(i,j)^{-}} \tag{3.1}$$

where

$$Cov(i,j)^{+} = Cov(i,j), \forall i \geq E(i) \; or \; j \geq E(j)$$

$$Cov(i,j)^{-} = Cov(i,j), \forall i < E(i) \; and \; j < E(j)$$

Since the most important objective of selecting portfolio is to reduce risk, we have to choose shares which can also reduce risk when necessary. In other words, if one of the shares faces huge devaluation, the other shares should be capable of compensating for that. If we look at the history of the shares relative to each other, they can have three different states:

1. Both shares have simultaneously been higher than their average yield.
2. One of the shares has been higher than its average yield while the other one has been lower than its average yield.
3. Both shares have simultaneously been lower than their average yield.

Therefore, we use covariance since it accurately depicts the exact extent of these three states. Positive covariance indicates states 1 and 2 and negative covariance indicates state 3.

Proposed Model:

1. At first, we calculate the initial value of the pheromone function using Equation 3.2 and the heuristic function is calculated from Equation 3.3.

   $$\Delta\tau_{ij}(t, t+1) = \sum_{k=1}^{m} \Delta{\tau^{k}}_{ij}(t, t+1) \tag{3.2}$$

   $$\Delta{\tau^{k}}_{ij}(t, t+1) = \frac{Q}{d_{ij}} \tag{3.3}$$

   In Equation 3.2, the value of the pheromone for ant K on the edge d_{ij} is calculated and in Equation 3.3 the entire pheromone on that edge after passing of m ants is calculated.

2. Determining the probability function for selecting the next city (firm's shares) using the following formula:

   $$\frac{\zeta_{ij}.\eta_{ij}}{\sum \zeta_{ij}.\eta_{ij}} \tag{3.4}$$

   Equation 3.4 is the probability function for selecting the next city (firm's shares) and it is calculated for all the cities where an ant can pass through. Finally, based on the determined weight, one of the next cities is selected for each simulation.

   ζ is the pheromone present in the edge between the two firms and η indicates the efficiency of the firm.

3. Randomly determining the next path relative to the probability of referring to each firm.

4. Updating the pheromone value in paths between firms.

   $$\zeta_{ij}(\mathrm{t}+1) = (1-\rho).\zeta_{ij}(\mathrm{t}) + \Delta\zeta_{ij} \tag{3.5}$$

After selecting the next city and before selecting the city after that, we have to update the pheromone value in all the paths. This update has two steps. The first step is the reduction of available pheromones due to pheromone evaporation. The evaporation rate is denoted by ρ. The second step involves adding the pheromone of the ant which passed through the last path and in this step, only this very path will be updated.

Updating is carried out by assuming that in step m, i.e. before the entrance of ant m, and based on the pheromone on the edges, by assuming the presence of n firms, we will have matrix A.

$$A = \begin{bmatrix} P_{11}\, P_{12} \;\ldots .\, P_{1j} \;\ldots .\, P_{1n} \\ \\ P_{j1} \;\ldots .\, P_{ij} \;\ldots .\, P_{jmn} \\ \\ P_{n1}\, P_{n2} \;\ldots .\, P_{nj} \;\ldots .\, P_{nn} \end{bmatrix}, \sum_{i=1}^{n} P_{ij} = 1, \forall i = 1, \ldots, n$$

After passing of ant $m + 1$, this matrix will change into matrix B.

$$B = \begin{bmatrix} \lambda_{11}P_{11} & \lambda_{12}P_{12} & \ldots & \lambda_{1n}P_{1n} \\ \\ \lambda_{j1}P_{j1} & \lambda_{j2}P_{j2} & \ldots & \lambda_{jn}P_{jn} \\ \\ \lambda_{n1}P_{n1} & \lambda_{n2}P_{n2} & \ldots & \lambda_{nn}P_{nn} \end{bmatrix},$$

In order to normalize the data, the matrix depicting the pheromone value for the passing of ant $m + 1$, is defined as follows:

$$C = \begin{bmatrix} \zeta_{11}\, \zeta_{12} \;\ldots .\, \zeta_{1j} \;\ldots .\, \zeta_{1n} \\ \\ \zeta_{j1} \;\ldots .\, \zeta_{ij} \cdot \zeta_{jmn} \\ \\ \zeta_{n1}\, \zeta_{n2} \;\ldots .\, \zeta_{nj} \;\ldots .\, \zeta_{nn} \end{bmatrix}, \sum_{i=1}^{n} \zeta_{ij} = 1, \forall i = 1, \ldots, n$$

where $\zeta_{ij} = \frac{\lambda_{ij}P_{ij}}{\sum_{j=1}^{n} \lambda_{ij}P_{ij}}$.

5. Four factors including Price-to-Earnings ratio(P/E), annual returns, ROA, and Earning Per Share(EPS) are selected as indicators (variables) and the matrix is developed.

$$dis(i, j) = \frac{av(j) \times Risk(j|i)}{av \times var(j)} \tag{3.6}$$

where $dis(i, j)$ indicates the efficiency of firm j compared to firm i and we used the yield of firm j ($av(j)$), average yield of all firms (av), risk present in firm j ($var(j)$) and the risk due to simultaneous selection of both firms ($cov(i, j)$) for calculating it.

6. At the end of each run of the program and after the passage of all the ants, the average thickness of a firm's edges will indicate the importance of the firm, using

which we can determine the probability functions of the firms relative to different factors.

$$\begin{bmatrix} P_{11}\ P_{12}\ \ldots.P_{1j}\ \ldots.P_{1m} \\ P_{j1}\ \ldots.P_{ij}\ \ldots.P_{jm} \\ P_{n1}\ P_{n2}\ \ldots.P_{nj}\ \ldots.P_{nm} \end{bmatrix}$$

where m is the number of selected factors and n is the number of selected firms. However, in order to carry out the normalization, the sum of the probabilities in each column must be equal to one.

After that, for all the paths proposed in the above iterations, the entropy is calculated. Finally, the path with the lowest entropy is selected as the best path and the corresponding portfolio will be selected as the best portfolio.

$$Entropy(j) = -\sum_{i=1}^{n} P_{ij} log P_{ij}\ j = 1,2, \ldots, m \tag{3.7}$$

$$Total\ Entropy(j) = \prod_{j=1}^{m} \text{Entropy(j)} \tag{3.8}$$

4. CASE STUDY REVIEW

Of all companies listed in Tehran Stock Exchange, 55 firms which were active during the period from 2003 to 2013 with high returns and low risks were selected as the sample of the study and their data were analyzed. Then, the tables related to the information on four criteria including price to earnings ratio, annual yield, ROA, and EPS were entered into the algorithm coded into MATLAB software application. After determining the number of variables, the number of ants, range of variables' changes, and entropy indicators, the ants started their job from a firm and determined a value for it based on the extent of pheromone left by previous ants. This process was repeated for all the ants in the colony and in each stage, the best ant was selected and the corresponding entropy was calculated. Finally, the portfolio with the lowest entropy was selected as the output of the software application. Moreover, this application was also assisted by the average function and the competency function for each ant.

5. RESULTS

In this study, we sought to propose a model by combining ant colony algorithm and entropy to predict and explain the stock market in order to help investors operating in this field, so the analysis and coding of the model was carried out accordingly. We also defined a new covariance to be used in financial problems. In this study, to select the portfolio, we used a non-linear model based on ant colony algorithm and entropy. Accordingly, at first we developed the possible portfolios of the selected firms and then, based on the ant colony algorithm and the indicator matrix, the portfolio with the lowest entropy was chosen by simulating the movement of ants in MATLAB software application.

REFERENCES

1. Chang, P.T. and Lee, J.H. (2012). A Fuzzy DEA & Knapsack formulation integrated model for project selection. *Computers & Operations Research*, 39, 112-125.
2. Chiam, S.C., Tan, K.C. and Al-Mamun, A. (2008). A memetic model of evolutionary PSO for computiational finance applications. *Experts systems with Applications*, 36, 3695-3711.
3. Dorigo, M., Birattari, M. and Stutzle, T. (2006). Ant Colony Optimization. *IEEE Computational Intelligence Magazine*, 1(4), 28-39.
4. Dorigo, M. and Gambardella, L.M. (1997). Ant Colony System: A Cooperative Learning Approach to the Traveling Salesman Problem. *IEEE Transactions on Evolutionary Computation*, 1, 53-66.
5. Dorigo, M., Maniezzo, V. and Colorni, A. (1996). The Ant System: Optimization by a Colony of Cooperating Agents. *IEEE Transactions on Systems, Man, and Cybernetics*, Part B, 26, 29-41.
6. Farnoosh, R., Rahimi, M. and Kumar, P. (2016). Removing noise in a digital image using a new entropy method based on intuitionistic fuzzy sets In *Proceedings of the International Conference on Fuzzy Systems,* Vancouver, BC, Canada, 24-29.
7. Lazo, J.G.L., Pacheco, M.A.C. and Vellasco, M.M.R. (2002). Portfolio selection and management using a hybrid intelligent and statistical system. In *Genetic Algorithms and Genetic Programming in Computational Finance* (pp. 221-238). Springer US.
8. Markechová, D. and Riečan, B. (2016). Logical Entropy of Fuzzy Dynamical Systems. *Entropy*, 18, 157.
9. Markechová, D. and Riečan, B. (2016). Entropy of Fuzzy Partitions and Entropy of Fuzzy Dynamical Systems. *Entropy*, 18, 19.
10. Markowitz, H. (1952). Portfolio Selection. *Journal of Finance*, 7, 77-91.
11. Najafi Moghadam, A., Roodpooshti, Fraydoon, R. and Farrokhi, M. (2014). Optimization of Stock Portfolio based of Ant Colony & Greay Theory. *International Research Journal of Applied and Basic Sciences*, 8, 780-788.
12. Sang, M.L. and Lerro, A.J. (1973). Optimization of Portfolio Selection for Mutual Funds. *Journal of Finance*, 28, 1087-1101.
13. Yan, W., Miao, R. and Li, S. (2007). Multi-period semi-variance portfolio selection: Model & numerical. *Applied Mathematics and Computation*, 194, 128-134.
14. Yari, G., Rahimi, M. and Kumar, P. (2017). Multi-period Multi-criteria (MPMC) Valuation of American Options Based on Entropy Optimization Principles. *Iranian Journal of Science and Technology, Transactions A: Science*, 41(1), 81-86.
15. Yari, G., Rahimi, M., Moomivand, B. and Kumar, P. (2016). Credibility Based Fuzzy Entropy Measure. *Australian Journal of Mathematical Analysis and Applications*, 13(1), 1-7.