Stock portfolio optimization using priority index and genetic algorithm

To cite this article: V D Vasiani et al 2020 J. Phys.: Conf. Ser. 1442 012031

View the article online for updates and enhancements.
Stock portfolio optimization using priority index and genetic algorithm

V D Vasiani, B D Handari and G F Hertono
Department of Mathematics, Faculty of Mathematics and Natural Sciences (FMIPA), Universitas Indonesia, Kampus UI Depok, Depok 16424, Indonesia

Corresponding author’s e-mail: bevina@sci.ui.ac.id

Abstract. Stock portfolio is a kind of investment which consists of several stocks. The aim of a stock portfolio is to minimize the risk of an investment and maximize the return on investment. To construct the optimum portfolio of stocks, one needs a strategy of stock selection and must determine the percentage of investment in each stock selected. In this paper, both the priority index method and genetic algorithm are applied to optimize the stock portfolios in terms of the return. Priority index is used in stock selection based on some parameters: price/earnings (P/E), earnings/share (EPS), wealth creation, undervaluation, and price per earnings/growth (PEG). Stock selection in each sector is determined by choosing the stocks which have a priority index score at least equal to the minimum priority index score of the selected stocks. The minimum priority index score of the selected stock is determined by using a certain scale parameter. The percentage of investment in each selected stock is then determined by using a genetic algorithm. The results showed that increasing the value of scale parameters does not always increase the average return. Moreover, the stock selection with a wealth creation parameter has a higher average return than without a wealth creation parameter. Stock selection using daily data has a higher average return than annual data. The results also showed that the method has an optimum period of up to five months to make an investment decision.

Keywords: genetic algorithm, optimization, priority index, stock portfolio.

1. Introduction

Investment is the placement of money or funds to obtain additional or certain advantages over the money or funds. Besides providing advantages (return), an investment also has a risk that is borne by the investor. The higher the rate of return expected by an investor, the higher the level of risk that should be covered by the investor [1]. The level of risk at a particular rate of return expectations can be minimized by forming a portfolio.

Portfolio optimization theory was first developed by Harry Markowitz in 1952 and known as the model of Markowitz Mean Variance [2]. Afterwards, Markowitz Mean Variance models is developed into other portfolio optimization models, such as the model of Capital Asset Pricing (CAPM) by William Shrape, CAPM taxes by Robert Hamada, quadratic programming, genetic algorithms, memetic algorithms, particle swarm optimization, differential evolution, and Tabu search [3].

In portfolio optimization, there are two important parts that affect the outcome of optimization: assets elections and the determination of the percentage of investments in each of the assets selected. In 2015, Sinha and his colleagues [4] completed a stock portfolio optimization problem using priority index and genetic algorithms. Priority index is used as basis for selection of constituent stocks portfolio and a genetic algorithm is used to determine the percentage of investments in each stock selected.

This paper conducts further analysis of the method described by Sinha et al. [4] to determine the role of the scale value in stock options to the average return generated by the portfolio, the role of the
parameters in the selection of stocks on average return generated by the portfolio, the role of the data types used in the selection of stocks on average return generated by the portfolio, and the optimal period of application of the method priority index and genetic algorithms in the optimization of stock portfolios.

2. Theoretical overview

In this section we will discuss the priority index calculation for stock selection and genetic algorithms to determine the percentage of investment. The priority index calculation refers to Sinha et al. [4] and genetic algorithm refers to Sivanandam and Deepa [5]. Priority index is the result of stock performance assessment by considering these five parameters: \( \frac{\text{price}}{\text{earning}} \) (P/E), \( \frac{\text{earning}}{\text{share}} \) (EPS), wealth creation, undervaluation, and \( \frac{\text{price per earning}}{\text{growth}} \) (PEG). The first step in calculating the priority index of a stock \( (IP_t) \) is to calculate scores of each stock on every parameter \( (S_{ij}) \). Suppose \( N \) is the number of shares, \( S_{ij} \) is a score of the \( i \)-th stock on parameter \( j \), \( X_{ij} \) is a the \( j \)-th parameter value of stock \( i \), \( \max_{ij} \) is the maximum value of the \( j \)-th parameter of stock \( i \), \( \min_{ij} \) is a minimum value of the \( j \)-th parameter of stock \( i \), with \( i = 1, 2, ..., N \), and \( j = 1,2,3,4,5 \). The scores of each stock on EPS and wealth creation respectively can be calculated by the following equation:

\[
S_{ij} = \frac{100(X_{ij} - \min_{ij})}{(\max_{ij} - \min_{ij})} \tag{1}
\]

The scores of stock on parameter P/E, undervaluation, and PEG follow this following equation:

\[
S_{ij} = \frac{100(X_{ij} - \min_{ij})}{(\max_{ij} - \min_{ij})} \tag{2}
\]

The priority index of a stock is the sum of the score on each parameter:

\[
IP_t = S_{i1} + S_{i2} + S_{i3} + S_{i4} + S_{i5} \tag{3}
\]

where \( IP_t \) is and priority index of the \( i \)-th stock and \( S_{ij} \) is the scores of the \( i \)-th stock on parameter \( j \).

A portfolio consists of several sectors, and each sector consist of several stocks. Based on equation (3), one can determine the stocks, which has the maximum or minimum priority indexes in each sector. Suppose \( s \) is the given scale value which is equal for all sectors. The maximum/minimum in equation (4) is the maximum/minimum priority index in the sector, and the minimum limit of priority index can be determined by:

\[
\text{Minimum limit} = \text{minimum} + s\ (\text{maximum} - \text{minimum}) \tag{4}
\]

The next step is to select stocks in each sector which have priority index that is greater or equal to the minimum limit on the corresponding sector. The selected stocks compose the portfolio.

The percentage of each stock selected is determined by a genetic algorithm. Genetic algorithm is based on the principles of genetics and natural selection. Suppose that there are \( n \) selected stocks, \( \mu_i \) is the return of stock \( i \), \( w_i \) is the percentage of stocks \( i \), and \( cov_{ik} \) is the covariance between stock returns \( i \) and stock return \( k \). The objective function:

\[
\max \sum_{i=1}^{n} \mu_iw_i \tag{5}
\]

\[
\min \sum_{i=1}^{n} \sum_{k=1}^{n} cov_{ik}w_iw_k \tag{6}
\]

with constrains:

\[
\sum_{i=1}^{n} w_i = 1 \tag{7}
\]

\[
w_i \geq 0 \tag{8}
\]
Fitness value is defined as [4]:

\[
\text{fitness} = \frac{\text{portfolio return}}{\text{portfolio standard deviation}}
\]  

(9)

In a genetic algorithm, a portfolio represented as one chromosome, where one chromosome consists of some genes. Each chromosome represents the percentage of each selected stock in one portfolio, and each gene represents the percentage of each stock.

The initial population is a collection of chromosomes in first generation. The population size specifies the number of chromosomes in each population in every generation. There are two methods to build the initial population, those are randomized and heuristic methods. This paper uses a random method.

There are several methods of chromosomes selection to the reproduction pool, such as rank selection, steady-state selection, roulette selection, and tournament selection. Roulette selection, as used in this paper, is a method based on the fitness of chromosomes.

Crossover combines the two chromosomes in the reproduction pool to produce offspring chromosomes. There are several methods, such as single point crossover, two-point crossover, n-point crossover, merge crossover, uniform crossover, heuristic crossover, and arithmetic crossover, the last of which is used here.

There are several methods of gene mutations, to maintain the diversity of the chromosomes in the population, such as gen exchange, exchange line, replacement value of a gene with its inverse, and a shift in genes values that used in this paper. The process of initial population, selection, crossover, and mutation is referred to in Sivanandam and Deepa [5].

3. Implementation and results

The purpose is to obtain an optimal stock portfolio consisting of stocks that are listed in the Standard & Poor's 500 (S&P 500) in 2014. The portfolio performance evaluation is done by comparing the average return generated by portfolio with the average return generated by the market portfolio. The Stock selection based on the parameters values data from January until December 2014 which is taken from www.ychart.com and www.gurufocus.com. The percentage of investment is determined by the selected stock returns data from January until December, 2014. The performance of the portfolio is calculated by comparing the average return generated by a portfolio to the average return generated by the market portfolio. In this process, data of selected stock returns and market return from January 2014 to December 2015 are needed. The return calculation is using the closing price data that can be obtain from www.finance.yahoo.com.

A parameter in stock selections is the scale value. Let s represent the scale value used to determine the minimum limit of priority index, \( s \in \{0.76; 0.77; 0.78; 0.79; 0.8; 0.81; 0.82; 0.83; 0.84; 0.85; 0.86; 0.87; 0.88; 0.89; 0.9; 0.91; 0.92; 0.93; 0.94; 0.95; 0.96; 0.97; 0.98; 0.99; 1\} \). There are some parameters that are required in determining the percentage of investment in each selected stock using a genetic algorithm, such as population size, number of generations, the probability of crossover, and mutation probability. Based on the best experiments results, this paper uses a population size of 100. Another parameter values are as follows [6]: total generations: 1000; the probability of crossover (P_c): 0.6; the probability of mutation (P_m): 0.4. Each genetic algorithm implementation runs 30 times.

To determine the role of the scale value in stock selection to the average return generated by the portfolio, the implementation of priority index method and genetic algorithm on the optimization of the stock portfolio is performed 25 times. Table 1 displays the average return generated by the portfolio on each scale value. It shows that the increase of the scale value resulted less number of chosen stocks. Increase the scale value does not guarantee an increase in the average return generated by the portfolio, either on the lowest, average, or on the highest case. Based on the result on table 1, for the next implementation, the scale value used is 0.84, since this value gives the average return more than the market average return while still minimizing the risk.

To determine the role of wealth creation parameter in the selection of stocks on average return generated by the portfolio, the implementation of priority index method and genetic algorithm on the optimization of a portfolio of stocks done in two cases: with and without wealth creation parameters. Table 2 displays the average return generated by the portfolio.
Table 1. Average return and the number of selected stock with a different scale value.

| Scale | Number of Selected Stocks in Portfolio | Average Return of Portfolio |
|-------|--------------------------------------|----------------------------|
|       |                                      | Lowest (%) | Average (%) | Highest (%) |
| 0.76  | 65                                   | 6.69       | 7.88        | 9.24        |
| 0.77  | 60                                   | 7.43       | 8.74        | 10.15       |
| 0.78  | 55                                   | 9.16       | 10.23       | 12.04       |
| 0.79  | 52                                   | 10.08      | 10.99       | 12.56       |
| 0.8   | 47                                   | 8.50       | 9.61        | 11.06       |
| 0.81  | 44                                   | 8.27       | 9.41        | 11.33       |
| 0.82  | 38                                   | 9.25       | 10.36       | 11.72       |
| 0.83  | 35                                   | 7.73       | 9.62        | 11.77       |
| 0.84  | 34                                   | 8.96       | 11.42       | 12.25       |
| 0.85  | 29                                   | 9.15       | 10.30       | 12.28       |
| 0.86  | 24                                   | 8.95       | 10.49       | 12.04       |
| 0.87  | 22                                   | 9.07       | 10.42       | 11.87       |
| 0.88  | 21                                   | 10.78      | 12.32       | 13.67       |
| 0.89  | 20                                   | 9.84       | 11.95       | 14.46       |
| 0.9   | 19                                   | 10.41      | 11.89       | 13.93       |
| 0.91  | 16                                   | 10.45      | 11.72       | 13.85       |
| 0.92  | 15                                   | 10.10      | 11.75       | 15.00       |
| 0.93  | 14                                   | 9.98       | 11.99       | 15.34       |
| 0.94  | 13                                   | 11.41      | 13.50       | 16.67       |
| 0.95  | 11                                   | 12.27      | 14.89       | 17.83       |
| 0.96  | 11                                   | 13.08      | 15.20       | 19.00       |
| 0.97  | 10                                   | 14.90      | 17.03       | 21.19       |
| 0.98  | 10                                   | 14.91      | 16.80       | 20.64       |
| 0.99  | 10                                   | 14.02      | 16.84       | 21.55       |
| 1     | 10                                   | 15.15      | 17.18       | 22.19       |

Table 2. Average return of portfolio with and without wealth creation parameter.

| Parameters            | Average Return of Portfolio |
|-----------------------|----------------------------|
|                       | Lowest (%) | Average (%) | Highest (%) |
| Without wealth creation| 1.65       | 3.02        | 4.71        |
| With wealth creation  | 9.47       | 11.43       | 14.08       |

Table 2 shows that the average returns of portfolio without a wealth creation parameter is less than the average return generated by stock selection using the parameter in all criteria.

To determine the role of the data types used in the selection of stocks on average return generated by the portfolio, the implementation is done with the following result at table 3. It shows the average return generated by stock selection using annual data is less than the average return of stock selection using daily data.

Finally, to determine the optimal period of application of the priority index method and genetic algorithms in the optimization of stock portfolios, table 4 shows the average return generated by the portfolio and the average return generated by market portfolio for each period investment.
Table 3. Average return of portfolio by data typed used.

| Parameters  | Average Return of Portfolio |   |   |
|-------------|-----------------------------|---|---|
|             | Lowest (%)                  | Average (%) | Highest (%) |
| Daily data  | 10.43                       | 11.60       | 13.79       |
| Annual data | 1.51                        | 3.00        | 5.19        |

Table 4. Average return with different investment periods.

| Investment Period (month) | Average Return of Portfolio | Average Return of Lowest (%) | Average Return of Highest (%) |
|---------------------------|-----------------------------|-----------------------------|-------------------------------|
|                           | Lowest (%)                  | Average (%) | Highest (%) | Lowest (%) | Highest (%) |
| 1                         | -0.17                       | -0.16        | -0.14        | 3.01       | 3.20        |
| 2                         | 6.39                        | 6.39         | 6.41         | 2.22       | 2.41        |
| 3                         | 4.98                        | 4.99         | 5.00         | 0.44       | 0.64        |
| 4                         | 5.68                        | 5.69         | 5.71         | 1.32       | 1.52        |
| 5                         | 4.37                        | 4.38         | 4.40         | 2.35       | 2.55        |
| 6                         | 0.11                        | 0.12         | 0.13         | 0.20       | 0.40        |
| 7                         | -5.07                       | -5.04        | -5.03        | 2.42       | 2.62        |
| 8                         | -7.74                       | -7.70        | -7.69        | -4.20      | -4.40       |
| 9                         | -11.60                      | -11.56       | -11.54       | -6.74      | -6.94       |
| 10                        | -2.29                       | -2.27        | -2.24        | 0.99       | 1.19        |
| 11                        | -3.15                       | -3.12        | -3.09        | 1.04       | 1.24        |
| 12                        | -4.03                       | -4.01        | -3.99        | -0.73      | -0.93       |

A one- to five-month investment period produces a higher average return than the average return generated by the market portfolio. For the other investment periods, the portfolio produce an average return less than the average return generated by the market portfolio.

4. Conclusions

Increasing the scale value in the stock selection process does not guarantee an increase in the portfolio average return. The wealth creation parameter is essential for the stock selection process. The average return obtained by daily data is higher than the average return of the annual data. Based on the used data, the priority index value method and genetic algorithms give an average return which is optimal for making stock portfolio investments up to five months ahead.

References

[1] Jones C L 2010 The Intelligent Portfolio: Pracitacl Wisdom on Personal Investing from Financial Engines (Hoboken: John Wiley & Sons)
[2] Lin C M and Gen M 2007 Appl. Math. Sci. 1 201–10
[3] Arriaga-González J A and Valenzuela-Rendón M V 2013 Reinforcement learning method for portfolio optimal frontier search Advances in Soft Computing and Its Applications (Lecture Notes in Artificial Intelligent vol 8266) ed F Castro et al. (Berlin: Springer-Verlag) pp 245–55
[4] Sinha P, Chandwani A and Sinha T 2015 Int. J. Syst. Assur. Eng. Manag. 6 447–65
[5] Sivanandam S N and Deepa S N 2008 Introduction to Genetic Algorithms (Berlin: Springer-Verlag)
[6] Sinha P and Goyal L 2012 Algorithm for Construction of Portfolio of Stocks using Treynor’s Ratio (Munich: Munich University Library) Munich Personal RePEc Archive (MPRA) Paper No. 40134