Available online at https://journal.rescollacomm.com/index.php/ijbesd/index

International Journal of Business, Economics and Social Development

e-ISSN 2722-1156
p-ISSN 27722-1164

Vol. 5, No. 2, pp. 227-234, 2024

Optimization of Investment Portfolio Mean-Variance Model Using Genetic Algorithm

Raynita Syahla[1*], Dwi Susanti[2], Herlina Napitupulu[3]

[1] *Undergrad Program in Mathematics, Faculty of Mathematics and Sciences, Universitas Padjadjaran, Sumedang, Jawa Barat, Indonesia*

[2,3] *Departement of Mathematics, Faculty of Mathematics and Sciences, Universitas Padjadjaran, Sumedang, Jawa Barat, Indonesia*

**Corresponding author email: raynita20001@mail.unpad.ac.id*

Abstract

The optimization of investment portfolio is aimed at finding the optimal combination of each stock with the goal of maximizing returns while minimizing risk through diversification. However, the question is how much funds should be invested to achieve the minimum risk. One of the approaches that has proven effective in building an optimal investment portfolio is the Mean-Variance model. The aim of this research is to determine the weights of the optimal portfolio components with the minimum risk. The data used consists of stocks included in the LQ45 index for the period from February 2020 to July 2021. Based on the research results, there are five stocks that form the optimal portfolio, namely ADRO, AKRA, BBCA, CPIN, and EXCL. The allocated weights for each stock are ADRO 9.896%, AKRA 32.049%, BBCA 30.749%, CPIN 13.949%, and EXCL 13.357%. The optimal portfolio generated by the Genetic Algorithm method has a risk of 0.000472 and an expected return of 0.000492.

Keywords: Mean-Variance, Lagrange Multiplier, Genetic Algorithm.

1. Introduction

The capital market plays a crucial role in nation's economic sector, serving as a fundamental pillar. Market activities have grown over time, driven by the increasing needs of the population, leading to heightened investments awareness. Various industries and enterprises utilize the capital market to attract investments and strengthen their financial positions. Hasani (2022) states that investment is a long-term capital investment activity in one or more assets with the expectation of gaining profits in the future. Stocks are particularly popular financial instruments traded on the capital market, offering attractive returns through dividends and capital gains.

There are many stocks traded in the capital market. By November 2023, the Indonesia Stock Exchange (IDX) listed 901 stocks, showcasing the market's diversity. Among its 43 stock indices, the LQ45 index is notable for measuring the price performance of 45 highly liquid and well-capitalized stocks. The LQ45 index is often used as a basis for portfolio construction. In stock investments, investors seek suitable returns aligned with their invested funds, making capital market investment a preferred choice due to its potential for higher returns compared to real asset or money market investments.

However, investing in the capital market entails risks, with higher returns often associated with greater risks. Therefore, investors face uncertainties regarding their investment outcomes, highlighting the importance of accurately estimating risk involved. Diversification, spreading funds across various securities, offers a strategy to mitigate investment risks. Yet, determining the optimal allocation of funds to minimize risks remains challenging. Efficient portfolio analysis aims to select the best stocks for an optimal investment portfolio formation. The Mean-Variance model, developed by Harry Markowitz capable to optimize the risk and return of their portfolios by considering the correlation between assets.

In forming an optimal portfolio, there exists a heuristic method capable of finding solutions to optimization problems in complex search such as portfolio formation. Genetic Algorithms, in particular, represent a heuristic optimization technique inspired by natural evolutionary processes. Compared to other heuristics, Genetic Algorithms yield superior results for multi-objective optimization problems (Setiawan & Rosadi, 2019). In the context of portfolio formation, Genetic Algorithms assist investor in selecting optimal weights for portfolio construction.

Based on the explanation above, this research involves forming an optimal Mean-Variance portfolio using Genetic Algorithm. The aim is to determine the weights of the optimal portfolio components with the minimum risk.

2. Literature Review

2.1. Stocks

Stocks can be defined as a form of participation or ownership by an individual or entity in a limited company or individual (Darmadji & Fakhruddin, 2006).

1) Return of Stocks

$$R_{it} = \frac{P_{it} - P_{i(t-1)}}{P_{i(t-1)}} \tag{1}$$

R_{it} : return of stock i at time t
P_{it} : price of stock i at time t
$P_{i(t-1)}$: price of stock i at time $t-1$

2) Expected Return of Stocks

$$E(R_{it}) = \frac{\sum_{t=1}^{m} R_{it}}{m} \tag{2}$$

$E(R_{it})$: expected return of stock i
R_{it} : return of stock i at time t
m : number of observation periods

3) Risk of Stocks

$$s_i^2 = \sum_{t=1}^{m} \frac{[R_{it} - E(R_{it})]^2}{m-1} \tag{3}$$

s_i^2 : variance of stock i
R_{it} : return of stock i at time t
$E(R_{it})$: expected return of stock i at time t
m : number of observations

4) Covariance
Risk measurement can be expressed in the form of covariance, which is a measure indicating the tendency of movement of two variables. The equation is as follows (Hartono & Rohaeni, 2021):

$$\sigma_{i,j} = \sum_{t=1}^{m} \frac{\left[\left(R_{it} - E(R_{it})\right) \cdot \left(R_{jt} - E(R_{jt})\right)\right]}{m} \tag{4}$$

$\sigma_{i,j}$: covariance of stock i and j
R_{it} : return of stock i at time t
R_{jt} : return of stock j at time t
$E(R_{it})$: expected return of stock i
$E(R_{jt})$: expected return of stock j
m : number of observations

5) Correlation Coefficient
Although covariance can be used to express the direction of stock movement, covariance figures are sensitive to measurement units. Therefore, it is necessary to calculate the correlation coefficient. The equation is as follows (Hartono & Rohaeni, 2021):

$$r_{i,j} = \frac{\sigma_{i,j}}{\sigma_i \sigma_j} \tag{5}$$

$r_{i,j}$: correlation coefficient of stock i and j
$\sigma_{i,j}$: covariance of stock i and j
σ_i : standard deviation of stock i
σ_j : standard deviation of stock j

2.2. Stock Portfolio

A stock portfolio is a collection of various stocks, owned by individuals or companies with different weights.

1) Expected Return of Portfolio

$$\mu_P = \sum_{i=1}^{n} w_i E(R_i) \tag{6}$$

μ_p : expected return of the portfolio
w_i : weight of stock *i*
$E(R_i)$: expected return of stock *i*
n : number of stocks in the portfolio

2) Risk of Portfolio

$$\sigma_P^2 = \sum_{i=1}^{n} w_i^2 \sigma_i^2 + \sum_{i=1}^{n} \sum_{j=1}^{n} w_i w_j \sigma_{ij} \tag{7}$$

σ_p^2 : risk of the portfolio
w_i : weight of stock *i*
w_j : weight of stock *j*
σ_i^2 : variance of stock *i*
σ_{ij} : covariance of returns stock *i* and stock *j*
n : number of stocks in the portfolio

2.3. Mean-Variance Portfolio Model

Modern portfolio theory was first introduced by Harry Markowitz in 1952. This theory states that return and risk are important factors to consider when forming a portfolio. The objective function used to determine the optimal portfolio is written as follows (Yunita, 2018):

$$\begin{aligned} min:&\ \textstyle\sum_{i=1}^{n} w_i^2 \sigma_i^2 + \sum_{i=1}^{n} \sum_{j=1}^{n} w_i w_j \sigma_{ij} \\ \text{s.t.}&\ \textstyle\sum_{i=1}^{n} w_i = 1 \\ &\ w_i \geq 0 \\ &\ R_p = \textstyle\sum_{i=1}^{n} w_i R_i \end{aligned} \tag{8}$$

2.4. Genetic Algorithm

The Genetic Algorithm was first introduced by John Holland in 1960 and further popularized by David Goldberg in 1980 (Haupt & Haupt, 2004). Genetic Algorithm are algorithms of numerical optimization inspired by natural selection and natural genetics according to Darwin's theory of evolution (Zainuddin & Samad, 2020). The Genetic Algorithm is an optimization technique that can tackle many difficult problems that conventional techniques cannot handle (Fanggidae & Lado, 2015). There are important definitions to be aware of in Genetic Algorithms (Satriyanto, 2009):

a) Gene, a value that represents basic units forming a specific meaning.
b) Chromosome, a combination of genes.
c) Individual, representing a single value or state that represents one of the possible solutions.
d) Population, a set of individuals that will be processed together in one evolutionary cycle.
e) Generation, representing one evolutionary cycle.

2.5. Investment Portfolio Optimization of Mean-Variance Model with Genetic Algorithm

Here are the steps of optimization using Genetic Algorithm:

1) Population initialization
   Population initialization involves determining the representation of genes and chromosomes within the population. The initial population is formed by randomly generating chromosomes according to the population size.

2) Evaluation

There are stages in evaluating the fitness value of an individual as follows:

a. Determining the objective function

The objective function represents the goal based on constrain to be sought for its solution. Based on equation (8), the objective function is:

$$\begin{aligned} min:\ & \sum_{i=1}^{n} \frac{s_i}{\sum_{i=1}^{n} s_i} \sigma_i^2 + \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{s_i}{\sum_{i=1}^{n} s_i} \frac{s_j}{\sum_{j=1}^{n} s_j} \sigma_{i,j} \\ s.t.\ & \sum_{i=1}^{n} \frac{s_i}{\sum_{i=1}^{n} s_i} = 1 \\ & \frac{s_i}{\sum_{i=1}^{n} s_i} \geq 0 \\ & R_p = \sum_{i=1}^{n} \frac{s_i}{\sum_{i=1}^{n} s_i} R_i \end{aligned} \tag{9}$$

b. Calculating the fitness value

The fitness value is a measure indicating the value of a function to achieve the goal. The fitness value is calculated using the following equation:

$$Fitness[i] = \frac{1}{1 + F_{obj}[i]} \tag{10}$$

3) Selection

Selection is the process of selecting individuals that will remain in a population. The selected individuals in the selection process will become parents in the next generation. The method commonly used is the roulette wheel method. In this method, the higher the fitness value of an individual, the greater the chance of being selected as a parent (Widodo, 2020).

There are several steps in the roulette wheel selection:

a. Calculate the probability of each chromosome being selected as follows:

$$P[i] = \frac{Fitness[i]}{\sum_{i=1}^{n} Fitness[i]} \tag{11}$$

b. Calculate the cumulative probability value as follows:

$$Q[i] = \sum_{i=1}^{n} P[i] \tag{12}$$

c. Generate random numbers $0 < R[i] < 1$ for each chromosome. If $Q[i] < R[i] < Q[i+1]$, then $Q[i+1]$ is selected as the new chromosome.

d. Select the parent chromosome for crossover based on crossover rate (ρc).

4) Crossover

Crossover involves crossing two chromosomes to produce new individuals inheriting basic characteristics from their parents. In the case of real number encoding, the crossover technique used is arithmetic recombination. The equation is as follows (Sandy et al., 2014):

$$\left.\begin{aligned} Child_1 &= \alpha \cdot Parent_1 + (1-\alpha) \cdot Parent_2 \\ Child_2 &= \alpha \cdot Parent_2 + (1-\alpha) \cdot Parent_1 \end{aligned}\right\} \tag{13}$$

Where α is a parameter controlling how much each parent contributes to the offspring's gene values.

5) Mutation

Mutation is the process of changing one or more genes within the same individual to produce a new individual. The number of chromosomes undergoing mutation is determined by mutation rate (ρm).

There are several steps in the mutation process:

a. Calculate the total length of genes in the population

b. Calculate the total genes in the population to undergo mutation as follows:

$$mutation\ count = \rho m \times total\ genes \tag{14}$$

c. Generate random numbers equal to the mutation count. The random numbers generated represent the location of the genes to be mutated.

d. Generate random numbers $[0, 1]$ equal to the mutation count. The random numbers generated are added to the corresponding genes.

6) Termination criteria
The termination criteria function to determine when the Genetic Algorithm will stop. The Genetic Algorithm will stop if it meets population convergence percentage (pk) or the counter generation (cg) value is greater than the total generation (jg).

3. Materials and Methods

3.1. Materials

In this study, the object used is daily stock closing proce from ten companies that are members of the LQ45 index. The data, sourced from yahoo finance with the period February 2020-July 2021. Data processing use Microsoft Excel tools and the Python programming language.

3.2. Methods

Here are steps in forming an optimal Mean-Variance portfolio using Genetic Algorithm:
Step 1. Calculating stock returns using equation (1) and expected stock returns using equation (2).
Step 2. Calculating the risk of stocks with positive expected returns using equation (3).
Step 3. Calculating stock covariance using equation (4) and forming them into the covariance matrix **Σ**.
Step 4. Calculating the correlation coefficient using equation (5).
Step 5. The objective function formulation from equation (9).
Step 6. Initialize the initial population by generating random values [0, 1] for 50 chromosomes.
Step 7. Calculating the objective function value for each chromosome.
Step 8. Initialize 50 population, $jg = 300$, and $cg = -1$
Step 9. Calculating the fitness value using equation (10).
Step 10. Calculating the population convergence (pk) with $cg = cg + 1$.
Step 11. Selection using roulette wheel method using equation (11) and (12) with $\rho c = 25\%$.
Step 12. Arithmetic recombination crossover using equation (12) with $\alpha = 0.5$.
Step 13. Mutation using equation (14) with $\rho m = 10\%$.
Step 14. Perform calculations until reaching the convergence criteria of $pk \geq 60\%$ or $jg \geq 300$.
Step 15. Calculating the expected portfolio return using equation (6) and portfolio risk using equation (7) for the Genetic Algorithm method. The portfolio with the smallest risk is considered optimal.
Step 16. The smallest risk value represent the most optimal portfolio.

4. Results and Discussion

4.1. Results

1) Stock price

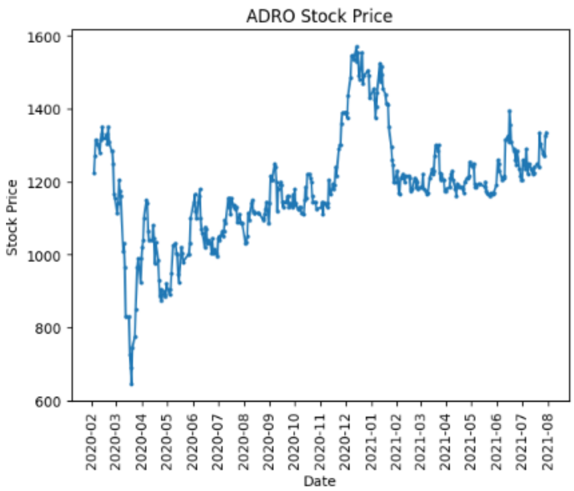

Figure 1: Closing stock price of ADRO

Based on Figure 1, it can be explained that the closing price of ADRO stock during the period from March 2020 to July 2021 experienced significant fluctuations. Firstly, there was a notable decrease from the beginning to the middle of March 2020, followed by a significant increase towards the middle of the month until April 2020. Subsequently, there was another increase in the middle of November 2020.

2) Expected return of stocks

The calculation results of the expected return of return using Microsoft Excel are presented in Table 1.

Table 1: Table of expected return of stocks

Stock Code	Expected Return
ADRO	0.000794
AKRA	0.000883
ASII	-0.000420
BBCA	0.000019
BBNI	-0.000647
CPIN	0.000470
EXCL	0.000443
ICBP	-0.000711
UNVR	-0.001459
WIKA	-0.001138

3) Risk of stocks

The risk calculation results of stocks with positive expected returns using Microsoft Excel are presented in Table 2.

Table 2: Table of risk of stocks

Stock Code	Risk of stocks
ADRO	0.001112
AKRA	0.001062
BBCA	0.000467
CPIN	0.001015
EXCL	0.001224

4) Covariance value of stock returns

The calculation results of stock covariances using Microsoft Excel are presented in Table 3.

Table 3: Table of covariance value of stock returns

	ADRO	AKRA	BBCA	CPIN	EXCL
ADRO	0.001112	0.000409	0.000351	0.000435	0.000582
AKRA	0.000409	0.001062	0.000255	0.000347	0.000419
BBCA	0.000351	0.000255	0.000467	0.000316	0.000347
CPIN	0.000435	0.000347	0.000316	0.001015	0.000474
EXCL	0.000582	0.000419	0.000347	0.000474	0.001224

5) Correlation value of stocks

Based on the covariance values, correlations between stocks were then calculated using Microsoft Excel, with the results shown in Table 4.

Table 4: Table of correlation value of stock returns

	ADRO	AKRA	BBCA	CPIN	EXCL
ADRO	1	0.376005	0.486511	0.409393	0.499181
AKRA	0.376005	1	0.362874	0.334447	0.36722
BBCA	0.486511	0.362874	1	0.459363	0.459095
CPIN	0.409393	0.334447	0.459363	1	0.424989
EXCL	0.499181	0.36722	0.459095	0.424989	1

6) Optimization using Genetic Algorithm

The Genetic Algorithm calculation using Python was conducted through several experiments, with the stopping criteria when the variance value convergence to 66%. To facilitate convergence calculation, the variance value was rounded to four decimal places. Thus, the best results obtained are as follows:

Table 5: Table of convergence of variance values

Variance values	Convergence
0.0004	6%
0.0005	66%
0.0006	28%

Several chromosomes were selected as candidates for forming optimal portfolios as follows:

Table 6: Table of optimization results using Genetic Algorithm

ADRO	AKRA	BBCA	CPIN	EXCL	σ_P^2	μ_P
0.104997	0.332071	0.205078	0.097755	0.260098	0.000531	0.000541
0.098960	0.320490	0.307490	0.139490	0.133570	0.000472	0.000492
0.168378	0.239148	0.222568	0.213148	0.156758	0.000494	0.000518
0.262315	0.274064	0.144551	0.244454	0.074617	0.000530	0.000601
0.111919	0.248228	0.218308	0.217538	0.204008	0.000500	0.000504
0.069590	0.376990	0.198840	0.239220	0.115360	0.000515	0.000555
0.115080	0.285560	0.189210	0.167780	0.242370	0.000521	0.000533
0.189860	0.229190	0.245800	0.165850	0.169300	0.000491	0.000510
0.257153	0.220005	0.290253	0.208325	0.024264	0.000480	0.000512
0.262315	0.274064	0.144551	0.244454	0.074617	0.000530	0.000601

Based on Table 6, the portfolio with the smallest variance has a value of 0.000472, with the following weights for the portfolio’s constituent stocks: ADRO 0.09896, AKRA 0.32049, BBCA 0.30749, CPIN 0.13949, and EXCL 0.133357. The return value generated by this portfolio is 0.000492.

4.2. Discussion

According to the theory presented in the study by Halim & Yulianti (2020), for optimizing the mean-variance investment portfolio using Lagrange multiplier, the optimal portfolio can be found using equation (9). The optimal portfolio is achieved when the value of tau is 0.42, with the following weights for each stock: ADRO 31.5432%, AKRA 43.2914%, BBCA 5.3582%, CPIN 18.5054%, and EXCL 1.3018%. The expected return and risk values obtained are 0.000726 and 0.000606, respectively.

5. Conclussion

The optimal portfolio consists of stocks from ADRO, AKRA, BBCA, CPIN, and EXCL with the following weights 9.896%, 32.049%, 30.749%, 13.949%, and 13.357%. The expected return and risk values of resulting of optimal portfolio are 0.000492 and 0.000472, respectively.

References

Darmadji, T., & Fakhruddin, H. (2006). *Pasar Modal di Indonesia.* Jakarta: Salemba Empat.

Fanggidae, A., & Lado, F. R. (2015). *Algoritma Genetika dan Penerapannya.* Yogyakarta: Teknosain.

Halim, N. A., & Yulianti, A. (2020). Markowitz Model Investment Portfolio Optimization: a Review Theory. *International Journal of Research in Community Service* , 14-18.

Hartono, N. P., & Rohaeni, O. (2021). Menentukan Portofolio Optimal Menggunakan Model Markowitz. *Jurnal Riset Matematika*, 57-64.

Hasani, M. N. (2022). Analisis Cryptocurrency Sebagai Alat Alternatif dalam Berinvestasi di Indonesia pada Mata Uang Digital Bitcoin. *Jurnal Ilmiah Ekonomi Bisnis*, 329-344.

Haupt, R. L., & Haupt, S. E. (2004). *Practical Genetic Algorithm, 2nd edition*. New Jersey: John Willey and Sons.

Sandy, I. A., Ichsan, M. H., & Setyawan, G. E. (2018). Analisis Performa Jaringan Sensor Nirkabel Berdasarkan Penentuan Lokasi Node yang Telah Diimplementasikan dengan Algoritma Genetika. *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, 3421-3431.

Satriyanto. (2009). *Algoritma Genetika*. Jakarta: Duta Ilmu Press.

Setiawan, E. P., & Rosadi, D. (2019). Model Pengoptimuman Portofolio Mean-Variance dan Perkembangan Praktisnya. *Jurnal Optimasi Sistem Industri*, 25-36.

Widodo, A. W., & Mahmudy, W. F. (2010). Penerapan Algoritma Genetika Pada Sistem Rekomendasi Wisata Kuliner. *Kursor, vol. 5 no. 4*. 205-211.

Yunita, I. (2018). Markowitz Model dalam Pembentukan Portofolio Optimal (Studi Kasus Pada Jakarta Islamic Index). *Jurnal Manajemen Indonesia*, Vol. 18 No. 1.

Zainuddin, F. A., & Samad, M. F. (2020). A Review of Crossover Methods and Problem Representation of Genetic Algorithm in Recent Engineering Applications. *International Journal of Advanced Science and Technology* , 759-769.