Optimal Portfolio Selection Using Mean Variance Model Based on Genetic Algorithm
An Empirical Study in a Sample of Algerian Stocks Exchange

Boudjenane Khaldia [1*]
[1] Department of Economics, Ibn Khaldoun University of Tiaret (Algeria)

Received: 30/10/2018 ; **Revised:** 07/11/2018 ; **Accepted:** 12/01/2019

Summary:

In this paper we tend to select an optimal portfolio from its different stochastic models. So the aim is to propose a new technique of optimization through the mean variance model based on genetic algorithm. This latter is used to minimize portfolio risk and maximize portfolio return of Algerian stocks exchange in order to prove the performance of the proposed technique which is implemented in three firms stocks.

The findings of this research allowed us to validate the performance of proposed technique which is strongly linked to the selected objective function.

Keywords: genetic algorithm; optimal portfolio; means variance; risk; expected return.

Jel Classification Codes : C06; G11

*Boudjenane Khaldia:E-Mail **: dehbias60@gmail.com**

I- Introduction :

The portfolio selection problem was initially developed by "Markowitz" in 1952. This problem is based on mathematical programming methods to find the optimal investment portfolio, which can maximize the portfolio return and minimize the portfolio risk at the same time. However; the risk and the return conflict with each other.

Markowitz mean variance model of stock portfolio selection and optimization is one of the well-known models in finance and it was the bedrock of modern portfolio theory.

In this model, it is assumed that asset returns follow a normal distribution. This involves the return on a portfolio of assets that can be perfectly characterized by the expected return and the associated risk measured by variance.

The model relies on several initial hypotheses regarding investors' behaviour on the market which can be cited as follows:

- Perfect and competitive markets: no tax. no transaction cost and the securities and assets are perfectly divisible;
- All investors are risk averse; all investors have the same beliefs;
- Security returns are jointly normal distribution;
- Dominant principle: an investor would prefer more return to less and would prefer less risk to more.

An assets' portfolio is efficient as long as no other portfolio with the same rate of return as the initial one; but with a lower risk can be created .The expected return variability is used to measure of the risk.

The purpose of this paper is to prove ***that portfolio selection problem can be successfully solved by mean variance based on genetic algorithm?***

The outline of this paper is organized as follows:

- **Section 2:** describe the portfolio optimization;
- **Section 3:** investigates basic structure of genetic algorithm;
- **Section 4:** dedicated to introduce our proposed algorithm was introduced;
- **Section 5:** provides our computational results using Matlab;

Finally; Section 6 presents our conclusions.

II- PORTFOLIO OPTIMIZATION:

Portfolio is to deal with the problem of how to allocate wealth among several assets. The portfolio optimization problems have been one of the important research fields in modern risk management. (Simona Dinu.Gabriela Andrei. 2012.p 3).

Generally speaking, an investor always prefers to have thv e return on his portfolio as large as possible. At the same time, he also wants to make the risk as small as possible. However a high return always accompanied with a higher risk. Markowitz introduced the mean variance model, which has been regarded as a quadratic programming problem.

2.1. Mean–variance model

Markowitz was the first to apply variance or standard deviation as a measure of risk. An n stocks portfolio rate of return is given by the weighted average of the average returns of the constitutive stocks. This average is between the limits of the best and the worst return of the portfolio stocks in terms of the weights of the constitutive stocks w_i. (Yang X.. 2006.p 11).

$$\mu_p = \sum_{i=1}^{n} w_i \cdot \mu_i \qquad (01)$$

Where:

- n: is the number of different invested assets.
- w_i:$i = 1. \dots . n$: is the weight for asset i.
- i: is the decision variable of the model.
- μ_p: is the return of portfolio.
- μ_i: is the expected return of asset.

The portfolio risk depends on three factors:

- the risk of every stock included in the portfolio;
- the covariance between the rates of return for assets in the portfolio;
- weightings of its constituent stocks.

Thus. the risk of a portfolio of n stocks is given by:

$$\sigma_p^2 = \sum_{i=1}^{n} \sum_{j=1}^{n} w_i \cdot w_j . \sigma_{ij} \qquad (02)$$

So

$$\sigma_p^2 = w' \sum w \qquad (03)$$

Where:

$$w = \begin{bmatrix} w_1 \\ \vdots \\ w_n \end{bmatrix} \qquad (04)$$

And

$$\sum = \sigma_{ij} = \begin{bmatrix} \sigma_n & \cdots & \sigma_{1n} \\ \vdots & \ddots & \vdots \\ \sigma_{n1} & \cdots & \sigma_{nn} \end{bmatrix} \quad \textbf{(05)}$$

The the variance of the total return in (01) is:

$$\mu_p = w'.\mu_i \quad \textbf{(06)}$$

Where we use matrix notation for the last equality w is a column vector whose components are the w_i ; w' is the row vector that is the transpose of w. and $\sum$ is the covariance matrix whose entries are the variances and covariances. (M. Gen. 2006. pp. 749–773).

$$\textbf{(07)}\sigma_{ij} = p_{ij}.\sigma_i.\sigma_j$$

Where:

- $\sum = \sigma_{ij} = cov(i.j)$: Covariance between stock i and stock j .
- $\sigma_i.\sigma_j$: Standard deviations of stock i and stock.
- p_{ij} : Correlation coefficient between stock i and stock j .

These quantities must be calculated statistically.

The covariance matrix derives from financial time series,which stock returns tracked over a certain period of time.

The identification of efficient portfolios is realized by determining the portfolio structure with a given rate of return and the lowest risk. The efficient layout of portfolio must determine a weighted average of the expected returns of stocks μ_i equal to the expected portfolio rate of return μ_p:

$$\sum_{i=1}^{n} w_i.\mu_i = \mu_p \quad \textbf{(08)}$$

Thus, the efficient frontier (the curve that joins all the best possible combinations of the n-stocks portfolios) is generated by the solutions of the following optimization problem:

$$min\sigma_p^2 = min\sum_{i=1}^{n}\sum_{j=1}^{n} w_i.w_j.\sigma_{ij} \quad \textbf{(09)}$$

$$\sum_{i=1}^{n} w_i.\mu_i = \mu_p \quad \textbf{(10)}$$

$$\sum_{i=1}^{n} w_i = 1.0 \leq w_i \leq 1.i = 1......n \quad \textbf{(11)}$$

The portfolios on the efficient frontier represent the set of Pareto-optimal portfolios; each portfolio on this frontier has the maximum expected return for a given amount of risk. or alternatively, the minimum risk for any given level of return.

In fact; the mathematical problem can be expressed in various ways:

- Minimize risk for a specified expected return ;
- Maximize the expected return for a specified risk;
- Minimize the risk and maximize the expected return. (H. Markowitz. 1952.p 34).

$$min\sigma_p^2 = min \sum_{i=1}^{n} \sum_{j=1}^{n} w_i \,.\, w_j . \sigma_{ij} \quad ; min\,\{risk(w_i)\} \qquad (12)$$

And

$$\max \mu_p \sum_{i=1}^{n} w_i \,.\, \mu_i \quad ; max\,\{return\,(w_i)\} \qquad (13)$$

Subject to:

$$\sum_{i=1}^{n} w_i = 1 \qquad (14)$$

A portfolio selection problem in the mean variance context can be written as:

$$\min \lambda \left[\sum_{i=1}^{n} \sum_{j=1}^{n} w_i \,.\, w_j . \sigma_{ij}\right] - (1-\lambda)\left[\sum_{i=1}^{n} w_i \,.\, \mu_i\right] \qquad (15)$$

Subject to:

$$\sum_{i=1}^{n} w_i = 1 \qquad (16)$$

This approach generates non-dominated solutions by varying the risk (λ) coefficient: from the minimum variance portfolio $(\lambda\ = 1)$ to the maximum return portfolio $(\ \lambda = 0)$.

Any value of l inside the interval (0. 1) represents a trade-off between the mean return and the variance; generating a solution between the two extremes (l = 0 and 1).

III– GENETIC ALGORITHM:

Based on the Darwin principle ‘‘the survival of the fittest” in nature; genetic algorithm GA was first initiated by Holland’s and has rapidly become the best-known evolutionary.

GA is stochastic search technique based on the mechanism of natural selection and natural genetics. (H. Markowitz. 1952.p 34).

The central theme of research on GA is to keep a balance between exploitation and exploration in its search to the optimal solution for survival in many different environments. Typically. Goldberg gave an interesting survey of some of the practical work carried out in this era and made clear of the general structure of GA.

Michalewicz (1996) did not restrict to the binary string encoding in Holland's GA and applied the GA to all possible encoding strategies to solve the practical optimization problems. GA has been theoretically and empirically proved to provide a robust search in complex search spaces. (M. Ehrgott. K. Klamroth. Schwehm. 2004. pp752 – 770). Many research papers and dissertations have established the validity of GA approach in function optimization problems and application problems Genetic Algorithm. differing from conventional search techniques, and starts with an initial set of random solutions. population. Each individual in the population is called a chromosome which representing a solution to the problem. The chromosomes evolve through successive iterations, called generations. (Kyong. Tae & Sungky. 2005.pp 371–379).

During each generation, the chromosomes are evaluated by taking some measures of fitness. To create the next generation with new chromosomes, called offspring. The offspring are formed by merging two chromosomes from current generation using the crossover operator and or modifying a chromosome using the mutation operator. A new generation is selected according to the fitness values of the parents and offspring, and then weeds out poor chromosomes so as to keep the population size constant. The algorithms converge to the best chromosome, which hopefully represents the optimum or suboptimal solution to the problem. (Srinivas . Lalit.1994.pp.18-20).

The basic steps in GA are shown as follows:

Step 1: Initialize a randomly generated population;

Step 2: Evaluate fitness of individual in the population;

Step 3: Apply elitist selection: carry on the best individuals to the next generation from reproduction, crossover and mutation;

Step 4: Replace the current population by the new population.

Step 5: If the termination condition is satisfied then stop, else go to Step 2.

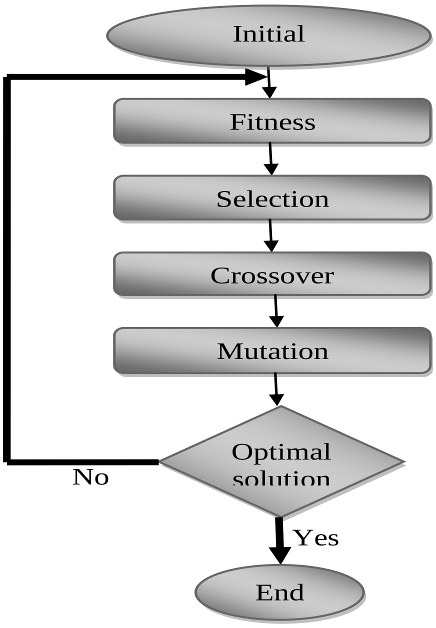
Through this reproduction once, the children of two chromosomes are generated. The reproduction process is operated until all chromosomes of a new population have been generated thoroughly;

Through specified maximum generations, the best solution ever found is the answer.(Tun-Jen Chang. Sang-Chin Yang. 2009. p 105).

IV- THE PROPOSED GA PORTFOLIO OPTIMIZATION

The proposed genetic algorithm for portfolio optimization problems based on the GA steps discussed in the previous section. This section we will describe in detail how to implement the proposed method. The detail procedure is illustrated as follows:

Figure (1)
flowchart of genetic algorithm.



4.1. Population initialization

This paper used a population of 100 subjects. Parents were chosen by binary tournament selection which works by forming two pools of individuals. each consisting of two individuals drawn from the population randomly. The individuals with the best fitness; one taken from each of the two tournament pools, are chosen to be parents.

4.2. Fitness objective function evaluation

To evaluate the fitness of the portfolio estimated and risk are used. We calculate them as follow:

$$\textit{Fitness function} = \lambda\left[\sum_{i=1}^{n}\sum_{j=1}^{n} w_i \,.\, w_j .\sigma_{ij}\right] - (1-\lambda)\left[\sum_{i=1}^{n} w_i \,.\, \mu_i\right] \quad (17)$$

Here λ represent the risk.

4.3. Reproduction, crossover,and mutation

In this section we will describe how the genetic operators are modified and how they performed in our algorithm. Children in our GA are generated by uniform crossover. In uniform crossover two parents have a single child.

V- RESEARCH RESULTS

The main objective of this paper is to illustrate via a Five (05) asset portfolio the efficiency of the GA in solving portfolio selection and optimisation problems. In order to achieve this goal the objective of the fitness function in the GA method is set as to maximize the return and minimize the risk of the portfolio.

We used historical daily data collected of the Algeria stock exchange. and we solved the problem using heuristic genetic algorithm and run on a personal computer with matlab.

The static stage of GA requires some parameters to attain a solution quality and sustain controllable evaluation of the process.

- Population size: $popSize = 100$;
- Maximum generation: $maxGen = 1000$;
- Crossover probability: $pC = 0.60$;
- Mutation probability: $pM = 0.30$.

In this numerical example. 05 stocks' data are used to demonstrate the proposed method.

- Stock Hotel Ourassi (AUR);
- Stock Alliance Assurance (ALL);
- Stock Rouiba (RUI);
- Stock Biopharm (BIO);
- Stock Saidal (SAI).

The portfolio average return and the portfolio variance are estimated using these historical data.

Table(1) the data

Period	t-4	t-3	t-2	t-1	t
Stock Hotel Ourassi	2.07	-0.14	-7.27	-3.13	1.24
Stock Alliance Assurance	5.4		-4.59	3.49	3.54
Stock Rouiba	0.13	-1.60	6.69	-5.07	3.46
Stock Biopharm	0.65	0.2	1.8	1.4	-0.9
Stock Saidal	-0.44	0.46	0.27	-1.5	1.7

Source: data from Algerian stocks exchange.

The mean return for each asset and the covariance matrix are given in the tables below:

Table(2) the mean returns for each asset

	AUR	ALL	RUI	BIO	SAI
Excepted stock returns	9.6	17.4	11.7	10.6	9.6

Source: results obtained using M.A.T.L.A.B

Table (3) the covariance matrix.

	AUR	ALL	RUI	BIO	SAI
AUR	**2.2**				
ALL	1.2	**1.5**			
RUI	-0.33	-0.39	**1.8**		
BIO	0.64	0.2	0.9	**1.5**	
SAI	-0.32	0.47	0.29	-0.4	**1.7**

Source: results obtained using M.A.T.L.A.B

From Table 2-3. The variance of Stock Rouiba refers to the variation in yields on average by a large estimated, which explains the increase in the size of the investment risks of the stock Biopharm caused by the low rate of return investigator and also the low volume of transactions. We also note that the risk of investing in stock Ourassi Hotel is great, but compared to Stock Rouiba is less.

We also note that the risk of stock Biopharm and stock of Alliance is high partly because of the continued decline in the price of mid-market and lower demand for stocks and the emergence of competition.

Table(4) Portfolio weights.

	w_1	w_2	w_3	w_4	w_5
Portfolio weights	0.01759	0.48034	0.34874	0.06735	0.07898

Source: results obtained using M.A.T.L.A.B

Sum of weigts =1

The expected return of a portfolio is the sum of the weight of each asset times the expected return of each asset. So the expected return of the portfolio is:

$$\mu_p = \sum_{i=1}^{n} w_i \cdot \mu_i$$

$$\mu_p = 0.0501$$

- Portfolio Return = 14.08912
- Portfolio Variance = 0.55007
- Portfolio Std.Dev = 0.75827

The variance of a portfolio of three assets can be expressed as:

$$\sigma_p^2 = \sum_{i=1}^{n} \sum_{j=1}^{n} w_i \cdot w_j \cdot \sigma_{ij}$$

$$\sigma_p^2 = [w_1 \cdot w_2 \cdot w_3 \cdot w_4 \cdot w_5] \left[\begin{bmatrix} \sigma_{11} & \cdots & \sigma_{15} \\ \vdots & \ddots & \vdots \\ \sigma_{51} & \cdots & \sigma_{55} \end{bmatrix} \right] \begin{bmatrix} w_1 \\ \vdots \\ w_5 \end{bmatrix}$$

$$\sigma_p^2 = 0.022357$$

And the standard deviation of the portfolio is:

$$\sigma_p^2 = \sqrt{\sigma_p^2} = \sqrt{0.022357} = 0.14952$$

We observe that the risk of a portfolio of three assets is less than the risk of each financial asset in the end however more financial assets is large in the portfolio the risk becomes less. so Diversification of assets reduces financial risks. So, ‘don't put all your eggs in one basket’.

VI- CONCLUSION:

In this paper, a genetic algorithm method was applied to solve the optimal portfolio selection. The method was applied on a simple example of three asset portfolio; the results obtained are interesting and confirm the efficiency of the genetic algorithm for its fast convergence towards the better solution and its interesting computing time.

So genetic algorithm is robust to solve the corresponding optimization problems because these no concave maximization problems are with a particular structure and cannot be efficiently solved by the existing traditional optimization methods in different risk measures. GA prominent advantage over other exact search methods is its flexibility and its ability to easily obtain a good solution to a problem where the other deterministic methods cannot achieve optimality in an easy manner.

Referrals and references:

H. Markowitz.(1952). Mean-Variance Analysis in Portfolio Choice and Capital Market. Basil Blackwell. New York. 1987.p 34.

Kyong. J. O.. Tae. Y. K.. & Sungky. M.. (2005). Using genetic algorithm to support portfolio optimization for index fund management. Expert Systems with Applications.Vol.28.pp 371–379.

M. Ehrgott. K. Klamroth and C. Schwehm. (2004). An MCDM approach to portfolio optimization. European Journal of Operational Research. pp752 - 770.

M. Gen and R. Cheng. (2000). Genetic Algorithms and Engineering Optimization. New York: John Wiley & Sons.p 67 .

Statistics. H. Pham (ed.). New York: Springer-Verlag. pp. 749–773.

M. Srinivas and M. P. Lalit.(1994). Genetic Algorithms: A Survey. IEEE Computer. vol.27. pp.18-20.

Simona Dinu.Gabriela Andrei. (2012). A New Multi-Objective Genetic Algorithm for Use in Investment Management. International Journal of Computer Applications. Vol.58.p 3.

Tun-Jen Chang. Sang-Chin Yang.(2009). portfolio optimization problems in different risk measures using genetic algorithm. Expert Systems with Applications.Vol 36.p 105.

Yang X.. (2006). Improving portfolio efficiency: a Genetic Algorithm Approach. Computational Economics. 28.p 11.