Applications of Genetic Algorithm to Portfolio Optimization with Practical Transaction Constraints

Chieh-Yow ChiangLin[1]

[1]Department of Finance, National Kaohsiung University of Applied Sciences

Abstract

The portfolio optimization model, initially proposed by Markowitz in 1952 and known as mean-variance model (MV model), is applied to find the optimized allocation among assets to get higher investment return and lower investment risk. However, the MV model did not consider some practical limitations of financial market, including: (1) transaction cost and (2) minimal transaction lots. While these constraints are not considered in the model, the practicability of the model will be restrained. But when they are included in the model, the model will become an NP hard problem, which cannot obtain global optimal solution by traditional mathematics programming techniques. In this research, besides proposing various models to include afore-mentioned consideration in the MV model, genetic algorithms are applied to solve these models. Empirical tests in the Taiwan stock market are provided to prove the applicability of the techniques.

Keywords: portfolio, mean-variance model, genetic algorithms.

1. Introduction

The portfolio selection problem, initially proposed by Markowitz in 1952, applies mathematical programming method to find the optimal investment portfolio which can maximize the portfolio return and minimize the portfolio risk at the same time. However, the mean variance model (MV model) proposed by Markowitz has not taken some realistic transaction problems in account, for example, minimal transaction lot and transaction cost. After these two problems are formulated in the MV model, the problem will turn to be a mixed integer mathematical problem, which is hard to find the global optimal solution by traditional linear programming method. In this paper, the genetic algorithm (GA) is applied to solve the extended MV model, which includes constraints of minimal transaction lot and transaction cost. Empirical test in the Taiwan stock market are conducted to show the capability of the technique.

2. Models Development

The mean variance model can be formulated as shown in Model (1), which is a quadratic programming (QP) problem.

$$\text{minimize} \quad \sigma_p^2 = \sum_{i=1}^{n}\sum_{j=1}^{n} w_i w_j \sigma_{ij} \tag{1}$$

subject to:

$$\sum_{i=1}^{n} w_i r_i = r_p ,$$

$$\sum_{i=1}^{n} w_i = 1 ,$$

Where, n represents the number of different invested assets; w_i represents the invested weight for asset *i*, which is the decision variable of the model; r_p represents the expected return of the portfolio; r_i represents the expected return of asset *i*; σ_p^2 represents the expected risk of the portfolio; σ_{ij} represents covariance between asset *i* and asset *j*.

Given aspired portfolio return, Model (1) can minimize the portfolio risk to find an efficient portfolio which will locate on the efficient frontier of the feasible solution space. Taking the portfolio expected return as another criterion in the objective function, Model (1) can be reformulated as shown in Model (2)

$$\max imize \quad (1-\lambda)\cdot\sum_{i=1}^{n} w_i r_i - \lambda\cdot\sum_{i=1}^{n}\sum_{j=1}^{n} w_i w_j \sigma_{ij} \tag{2}$$

subject to:

$$\sum_{i=1}^{n} w_i = 1 ,$$

$$w_i \ge 0 , \; i = 1,\ldots,n ,$$

Here, λ represent the risk aversion parameter, 0□ λ □1.

Based on Model (2), three extended model can be developed to consider the realistic transaction problem. The second model as shown in Model (3) take the minimal trasaction lot-size problem into account; The third model as shown in Model (4) take the transaction cost into account; The forth model as shown in Model (5) take the minimal trasaction lot-size problem and trasaction cost into account.

$$maximize \quad (1-\lambda)\cdot\sum_{i=1}^{n} w_i r_i - \lambda\cdot\sum_{i=1}^{n}\sum_{j=1}^{n} w_i w_j \sigma_{ij} \tag{3}$$

subject to

$$\sum_{i=1}^{n} w_i \le 1 \ ,$$

$$w_i \ge 0\ ,\ i = 1,2\ldots,n\ ,$$

$$w_i = x_i c_i / I\ ,$$

$$x_{i,t+1} \in N\ ,$$

$$\sum_{i=1}^{n} x_i c_i \le C$$

where, $c_i = N_i p_i$ represents the minimal lot-size of asset *i*; N_i represents the minimal transaction unit for asset *i*; p_i represents the price of asset *i*.

$$\max imize \quad (1-\lambda)\cdot\sum_{i=1}^{n} w_{i,t+1} r_{i,t+1} - \lambda\cdot\sum_{i=1}^{n}\sum_{j=1}^{n} w_{i,t+1} w_{j,t+1}\sigma_{ij,t+1} - \sum_{i=1}^{n} k_{i,t+1}\,|\,w_{i,t+1} - w_{i,t}\,| \tag{4}$$

subject to:

$$\sum_{i=1}^{n} w_i = 1 \ ,$$

$$w_i \ge 0\ ,\ i = 1,2\ldots,n\ ,$$

where $w_{i,t+1}$ represents the invested weight for asset *i* at time period t+1; $w_{i,t}$ represents the invested weight for asset *i* at time period t; $r_{i,t+1}$ represents the expected return for asset *i* at time period t+1; $\sigma_{ij,t+1}$ represents the covariance between asset *i* and asset *j* at time period t+1.

$$\max imize \quad (1-\lambda)\cdot\sum_{i=1}^{n} w_{i,t+1} r_{i,t+1} - \lambda\cdot\sum_{i=1}^{n}\sum_{j=1}^{n} w_{i,t+1} w_{j,t+1}\sigma_{ij,t+1} - \frac{\sum_{i=1}^{n} k\cdot|\,x_{i,t+1} - x_{i,t}\,|\cdot c_{i,t+1}}{C_{t+1}} \tag{5}$$

subject to:

$$\sum_{i=1}^{n} w_{i,t+1} \le 1 \ ,$$

$$w_{i,t+1} = x_i c_i / C_{t+1}\ ,$$

$$w_{i,t+1} \ge 0\ ,\ i = 1,2\ldots,n\ ,$$

$$x_{i,t+1} \in N\ ,$$

$$\sum_{i=1}^{n} x_{i,t+1}\, c_{i,t+1} \le C_{t+1}$$

Where, $x_{i,t}$ represents the invested unit for asset *i* at time period t; $x_{i,t+1}$ represents the invested unit for asset *i* at time period t+1; I_{t+1} represents the invested upper bound at time period t+1; $c_{i,t+1}$ represents the minimal lot-size of asset *i* at time period t+1.

3. Genetic Algorithm Design

For solving Model (3) ~ (5), the genetic algorithm proposed by Holland (1975) is applied to solve the mixed integer problems. Mechanism designs of the genetic algorithm for solving different models are described in this section.

3.1 Parameter setting

As Srinivas and Lalit (1994) stated that there is no perfect way for parameters setting in GA except for experimental training. After multiple tests, in this paper, the population crossover rate, mutation rate and generations are set to be 100, 1.0, 0.03 and 3000, respectively.

3.2 Encoding

In this research, Model (2) is solved by mathematical programming. As to Model (3) and Model (5), the encoding method is shown as follows: ($x_i \in N$)

x_1	x_2	x_3	x_4	x_5	....	x_{n-1}	x_n

The encoding method Model (4) is shown as follows: $w_i \in R$

w_1	w_2	w_3	w_4	w_5	....	w_{n-1}	w_n

3.3 Fitness function

Since the "Roulette-wheel" selection is adopted for selection, the output value of the fitness function should be greater than or equal to 0. The e^U function is selected as the fitness function where U represents the objective function of the models.

3.4 Initialization

100 chromosomes are created by random at the initial stage. Each chromosome keeps as many numbers of genes as those of invested assets.

For Model (4), the value of genes locate between 0 and 1. As to Model (3) and Model (5), since the lot-size constraint, the decision variable must be an integer. The gene value can be calculated by:

$$\text{Gene value of } i = f\left(\frac{r \text{ *Initial asset value}}{\text{Initial value of asset } i}\right)$$

where, *f(x)* will generate an integer smaller than *x*. Normalization is necessary for Model (3) ~ (5).

3.5 Selection

Roulette-wheel method is adopted for selection.

3.6 Crossover

Single-point crossover method is adopted for crossover process.

3.7 Mutation

The mutation rate is set to be 0.03 in this research.

4. Empirical Studies

The data from the Taiwan stock market is selected for empirical tests. For performance comparison among different proposed models, we select representative 42 assets in Taiwan stock market and find the efficient frontiers for various models

Empirical test results show that Model (3) can select stocks by lot-size. From the comparison between Model (2) and Model (3), we find that the Sharpe ratios of efficient portfolio, averaged 11.15953, in Model (2) are superior to those, averaged 10.1618, in Model (3). It is reasonable because the constraint of lot-size investment shrinks the feasible solution space. Efficient frontier comparison between Model (2) and Model (2) is shown in Figure 1.

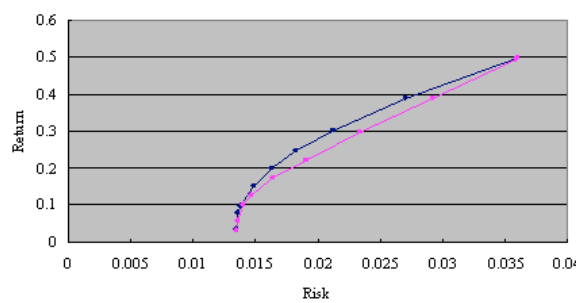

Fig. 1: Efficient frontier comparison between Model (2) and Model (3)

When comparing Model (2) to Model (4), efficient frontiers of two models are almost the same. The Sharpe ratios of efficient portfolio, averaged 22.66175, in Model (2) are still superior to those, averaged 22.20108, in Model (4) due to the consideration of transaction cost. Efficient frontier comparison between Model (2) and Model (2) is shown in Figure 2.

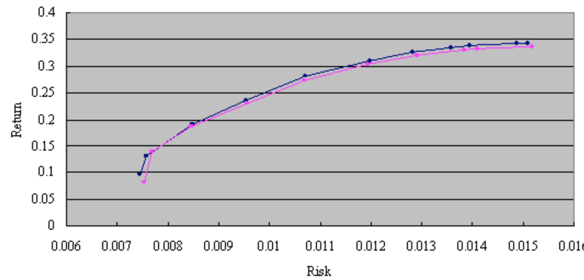

Fig. 2: Efficient frontier comparison between Model (2) and Model (4)

When comparing Model (2) to Model (5), it can be found that the Sharpe ratios of efficient portfolio in Model (2) are superior to those in Model (5) due to the consideration of lot-size investment and transaction cost.

References

[1] H. Markowitz, Portfolio selection, Journal of Finance 7, 1952, pp.77-91.

[2] J.H. Holland, Adaptation in Natural and Artificial Systems, Ann Arbor: The University of Michigan Press, 1975.

[3] M. Srinivas and M. P. Lalit, Genetic Algorithms: A Survey, IEEE Computer, vol.27, pp.18-20, 1994.