Portfolio Optimization Using an Integer Genetic Algorithm

Javier Fernández-Rendón
Facultad de Ciencias, UNAM
Mexico City, Mexico
ferjavi7@hotmail.com

Katya Rodríguez-Vázquez
IIMAS, UNAM
Mexico City, Mexico
katya.rodriguez@iimas.unam.mx

ABSTRACT

The portfolio selection is an essential component of fund administration because it contributes to economic growth of the investor. Many of the related works use the Markowitz's mean–variance portfolio selection model approach to solve this optimization problem. However, the use of continuous variables in this approach does not allow us to implement directly the obtained solutions because assets cannot be divided. This paper presents a portfolio selection model that involves integer variables, allowing a more realistic treatment. Due to the complexity of this mixed–integer nonlinear programming problem, a corresponding genetic algorithm is used to solve it.

Categories and Subject Descriptors

I.2.8 [**Artificial Intelligence**]: Problem Solving, Control Methods, and Search—*heuristic methods, genetic algorithms*; G.1.6 [**Numerical Analysis**]: Optimization—*constrained optimization, integer programming, nonlinear programming*; J.1 [**Administrative Data Processing**]: Financial

Keywords

Portfolio optimization, Markowitz model, Non–linear integer programming, Genetic algorithm

1. PROBLEM FORMULATION

The Markowitz's modern portfolio theory [3, 4] has made a new paradigm of portfolio selection problem in which the investor wants to form a portfolio with the lowest level of risk at a given desired expected return. This optimization problem consists in determining the optimal investing weight of each asset. However, this model assumes a perfect market in which assets can be divided infinitely and therefore can be traded in any fraction. Due to this strong assumption, this model cannot be implemented in practice and only gives an investment portfolio construction approximation. This approximation can be made with adjustments in the obtained solutions, commonly rounding this values. But this can lead to some issues. The first one is that previously obtained values of risk and expected return of the portfolio may have changed. And the second one is regarding the allocation of capital since it could exceed the available capital or it could stay well below this (a problem in stock markets, because there exist a minimum investment capital).

In order to handle these issues and look for more practical results, a model (based on Markowitz's model) which involves integer variables is here presented. This model takes into account the unit price of the assets and a minimum investment capital proportion to give a more realistic approach of portfolio selection. Let:

- n be the total number of assets available
- μ_i be the expected return of asset i $(i = 1, \ldots, n)$
- σ_{ij} be the covariance between assets i and j $(i = 1, \ldots, n;\ j = 1, \ldots, n)$
- r^* be the minimum desired expected return of investor
- C be the total available capital (monetary terms)
- p be the minimum proportion of capital to be invested
- c_i be the price of asset i (monetary terms)
- x_i be the number of asset i which model suggest to purchase (non–negative integer variable)

Then the proposed model is formulated as:

$$\min \sqrt{\sum_{i=1}^{n}\sum_{j=1}^{n} c_i\, x_i\, c_j\, x_j\, \sigma_{ij}} \tag{1}$$

$s.t.$

$$\sum_{i=1}^{n} c_i\, x_i\, \mu_i \geq r^*\, C \tag{2}$$

$$\sum_{i=1}^{n} c_i\, x_i \leq C \tag{3}$$

$$\sum_{i=1}^{n} c_i\, x_i \geq p\, C \tag{4}$$

$$x_i \geq 0 \text{ and integer}, \quad i = 1, \ldots, n \tag{5}$$

Eq. 1 is the objective function which determines the risk and it should be minimized. Eq. 2 ensures the desired minimum expected return of investor. Eq. 3 indicates the maximum capital allocation, whereas eq. 4 indicates the minimum capital allocation. Finally, eq. 5 indicates that x_i is a non–negative integer variable (i.e., no short sale is allowed).

GECCO'14, July 12–16, 2014, Vancouver, BC, Canada.
ACM 978-1-4503-2881-4/14/07.
http://dx.doi.org/10.1145/2598394.2602278.

2. GENETIC ALGORITHM DEFINITION

Since Genetic Algorithms (GAs) allow the treatment of very complex problems, such as mixed–integer nonlinear programming problems, a corresponding GA is proposed in order to deal with the portfolio optimization model here presented.

The design of the proposed GA is described as follows. The individuals are represented by an integer string values (as such are defined the decision variables); thus, each individual represents a portfolio investment where ith component indicates the number of ith asset to be bought. These values must be between 0 and $\lfloor \frac{C}{c_i} \rfloor$ in order to limit the search space.

For the evaluation of individuals the standard deviation as risk measure plus a penalty for each violated constraint are considered. This penalization is assigned depending on the amount of violation for each constraint of the problem and is expressed in terms of r^*, n and a penalty coefficient.

The selection process uses both a linear ranking [1] and a roulette wheel selection method [2] for prevents premature convergence. In addition, to preserve the fittest individual of the generation elitism is used.

As genetic operators, the intermediate recombination and the BGA mutation [5] are considered, both with rounding values for preserving representation of individuals. Also a repair strategy is implemented for those solutions that get out of the bounded search space assigning them the nearest limit value.

Finally, parameters are set as follows: a population size of 80 individuals, a total number of 10,000 generations, a crossover probability of 0.9 and a mutation probability of $1/n$ (the recommended for the BGA mutation method).

3. EMPIRICAL RESULTS

A portfolio optimization problem with a total number of 165 assets, with a sample historical data of 419 daily returns,[1] is formulated by proposed model considering an available capital of $100,000 and setting a 99% of this capital as a minimum required investment.

As it can see in Figure 1, both efficient frontiers one obtained by the defined GA and the other one obtained by means of rounding method to adjust the *MATLAB* generated solutions[2] are presented. Here the risk of *MATLAB* solutions is lower than the obtained by the defined GA; however the average investment of adjusted solutions is $98,619.87 (with a minimum of $93,662.26 and a maximum of $105,246.61) which implies the risk reduction but also infeasible solutions. On the other hand, the average investment of GA solutions is $99,977.54 (with a minimum of $99,753.02 and a maximum of $99,999.96); thus, they are feasible solutions with slightly higher risk.

4. CONCLUSIONS

In this paper, a portfolio selection model based on Markowitz's model is proposed. The presented model involved integer variables and a minimum capital investment constraint for a more realistic treatment of portfolio optimization. A corresponding GA is defined due to the proposed model complexity. The results show a good performance of the proposed GA.

Finally, in a future research, asset lots investment, buy–in threshold, cardinality and capitalization sector constraints with transaction costs and other risk measures will be considered.

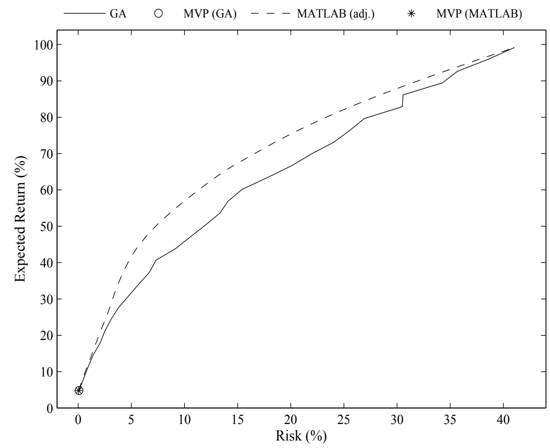

Figure 1: GA Efficient Frontier vs. MATLAB (adj.) Efficient Frontier.

5. REFERENCES

[1] J. E. Baker. Adaptive Selection Methods for Genetic Algorithms. In J. J. Grefenstette, editor, *Proceedings of the First International Conference on Genetic Algorithms and Their Applications*, pages 101–111, Hillsdale, NJ, USA, 1985. Lawrence Erlbaum Associates.

[2] D. E. Goldberg. *Genetic Algorithms in Search, Optimization and Machine Learning*. Addison-Wesley, Boston, MA, USA, 1989.

[3] H. M. Markowitz. Portfolio Selection. *The Journal of Finance*, 7(1):77–91, 1952.

[4] H. M. Markowitz. *Portfolio Selection: Efficient Diversification of Investments*. Yale University Press, USA, 1959.

[5] H. Mühlenbein and D. Schlierkamp-Voosen. Predictive Models for the Breeder Genetic Algorithm I. Continuous Parameter Optimization. *Evolutionary Computation*, 1(1):25–49, 1993.

[1]By convention, the standard deviations and expected returns have been annualized in order to obtain the annual volatility and the expected annual return of the portfolio.

[2]Using the Financial Toolbox.