Journal of Social and Economic Statistics

No.2, Vol.1, Winter 2012

A NEW GENETIC ALGORITHM TO SOLVE KNAPSACK PROBLEMS[*]

Derya TURFAN[a], Cagdas Hakan ALADAG[b], Ozgur YENIAY[c]

Abstract

Knapsack problem is a well-known class of optimization problems, which seeks to maximize the benefit of objects in a knapsack without exceeding its capacity. Various knapsack problems have been tried to be solved by decision makers a wide range of fields, the most important one of these is portfolio optimization. In recent few decades, heuristic methods have been widely used to solve hard optimization problems since they have proved their success in many real life applications. Genetic algorithm is one of these heuristic algorithms which have been successfully employed in a variety of continuous, discrete and combinatorial optimization problems. In this study, a new genetic algorithm is improved to solve a portfolio optimization problem efficiently.

Keywords: Genetic algorithms, Mean-variance optimization, Portfolio analysis, knapsack problem

JEL Classification: C44, C61, G11

Authors' Affiliation

[a] - Hacettepe University, Faculty of Science, Department of Statistics, Ankara, Turkey, e-mail: deryaturfan@hacettepe.edu.tr

[b] - Hacettepe University, Faculty of Science, Department of Statistics, Ankara, Turkey, e-mail: aladag@hacettepe.edu.tr (corresponding author)

[c] - Hacettepe University, Faculty of Science, Department of Statistics, Ankara, Turkey, e-mail: yeniay@hacettepe.edu.tr

---

*****An earlier version of this paper was presented at *The 6th International Conference on Applied Statistics*, November 2012, Bucharest.

1. Introduction

Knapsack problem has a central role in integer and nonlinear optimization, which has been intensively studied due to its immediate applications in many fields and theoretical reasons. In many applications, practical problems are often formulated as or approximated by continuous or integer KP. Examples of successful applications in last decades cover resource allocation (Bitran et al 1981; Bretthauer et al. 1997; Bretthauer et al. 2006), demand forecasting (Hua et al. 2006), portfolio selection (Li et al. 2006), network flows (Helgason et al. 1980; Ventura 1991), etc. (Zhang et al. 2012)

Gilli and Kéllezi (2000) has reviewed portfolio analysis problem as follows:

"The fundamental goal of an investor is to optimally allocate his investments among different assets. The pioneering work of (Markowitz 1952) introduced mean-variance optimization as a quantitative tool which carries out this allocation by considering the trade-off between risk (measured by the variance of the future asset returns) and return. Assuming the normality of the returns and quadratic investor's preferences allow the simplification of the problem in a relatively easy to solve quadratic program. Notwithstanding its popularity, this approach has also been subject to a lot of criticism. Alternative approaches attempt to conform the fundamental assumptions to reality by dismissing the normality hypothesis in order to account for the fat-tailedness and the asymmetry of the asset returns. Consequently, other measures of risk, such as Value at Risk (VaR), expected shortfall, mean absolute deviation, semi-variance and so on are used, leading to problems that cannot always be reduced to standard linear or quadratic programs. The resulting optimization problem often becomes quite complex as it exhibits multiple local extreme and discontinuities, in particular if we introduce constraints restricting the trading variables to integers, constraints on the holding size of assets, on the maximum number of different assets in the portfolio, etc. In such situations, classical optimization methods do not work efficiently and heuristic optimization techniques can be the only way out. They are relatively easy to implement and computationally attractive. The use of heuristic optimization techniques to portfolio selection has already been suggested by (Mansini et al. 1999), (Chang et al. 2000) and (Speranza 1996). This paper builds on work by (Dueck et al. 1992) who first applied a heuristic optimization technique, called Threshold Accepting, to portfolio choice problems. We show how this technique can be successfully employed to solve complex portfolio choice problems where risk is characterized by Value at Risk and Expected Shortfall."

In this study, the most active 10 stocks traded in Istanbul Stock Exchange 30 Index (ISE 30) was chosen and their proportion was calculated to obtain optimal portfolio with the proposed genetic algorithm. This paper gives in Section 2 formulation of standard Markowitz's mean-variance model and in section 3 brief information about the usage of genetic algorithms in portfolio optimization. The proposed genetic algorithm is introduced in Section 4. Section 5 presents the implementation. Finally, Section 6 provides the concluding remarks.

2. The portfolio selection problem

Portfolio optimization problem is a well-known problem in the literature. Various methods have been used to make decisions about assets (Egrioglu et al. 2012). There are different formulations for this problem. The general formulation of standard Markowitz's mean-variance model for the portfolio selection problem:

N: the number of assets available

μ_i : the expected return of asset i (i=1,…,N)

σ_{ij} : the covariance between assets i and j (i=1,…,N; j=1,…,N)

R^* : the desired expected return

w_i : the proportion ($0 \leq w_i \leq 1$) held of asset i (i=1,…,N)

Minimize $$\sum_{i=1}^{N}\sum_{j=1}^{N} w_i w_j \sigma_{ij} \quad (1)$$

subject to

$$\sum_{i=1}^{N} w_i \mu_i = R^* \quad (2)$$

$$\sum_{i=1}^{N} w_i = 1 \quad (3)$$

$$0 \leq w_i \leq 1, \quad i=1,\ldots, N \quad (4)$$

(1) minimizes the total variance (risk) associated with the portfolio whilst (2) ensures that the portfolio has an expected return of R^*. (3) ensures that the proportions add to one.

This formulation ((1)-(4)) is a simple quadratic programming problem for which computationally effective algorithms exist so there is (in practice) little difficulty in calculating the optimal solution for any particular data set. (Chang et al. 2000)

The Markowitz's mean-variance model to portfolio selection involves tracing out an efficient frontier. (Markowitz1959; Markowitz 1987)

A combination of assets, i.e. a portfolio, is referred to as "efficient" if it has the best possible expected level of return for its level of risk (usually proxies by the standard deviation of the portfolio's return) (Elton et al. 2011).

3. Genetic Algorithms

A genetic algorithm (GA) can be described as an “intelligent” probabilistic search algorithm. The theoretical foundations of Gas were originally developed by Holland (Holland 1975).

In genetic algorithms, an initial population containing constant number of chromosomes is generated randomly (regarding portfolio optimization, each chromosome represents the weight of an individual security) and an evaluation function is formed to evaluate the fitness of each chromosome, which defines if the chromosome represents a good solution. Using crossover, mutation and natural selection, the population will evolve towards a population that contains only the chromosomes with good fitness. The larger the fitness value is the better objective function the solution has. (Stancu et al 2010)

The basic steps in genetic algorithms are: (Chang et al. 2009)

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.

4. The proposed genetic algorithm

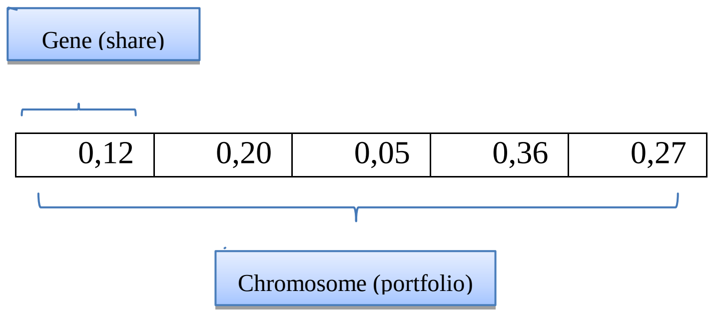
A new genetic algorithm is proposed to solve the problem defined in section 2. This section includes the basic elements and parameters of the proposed genetic algorithm.

<u>Solution</u>

Solutions are represented by chromosomes. The number of genes included in a chromosome equals to the number of shares. Each gene indicates the proportion ($0 \leq w_i \leq 1$) held of share i. (i=1,…,N)

There is an example of one solution in Figure 1.

Figure 1. An example of one solution



<u>Solution spaces</u>

The solutions satisfy constraints given at (2)-(5) compose of solution space of the portfolio optimization problem.

<u>The initial population</u>

To generate an initial population, a function coded in Matlab computer package was employed. This initial solution was generated depending on the desired expected return(R^*). All solutions included in initial populations satisfy the constraints of the related problem. The proposed genetic algorithm started from an initial solution.

<u>Fitness function</u>

Objective function of the portfolio optimization problem given in (1) was used as fitness function.

<u>Neighborhood structure</u>

To generate a new population, crossover operation was used in the proposed algorithm. A point was defined randomly and for two parents, the genes after this point changed places with each other.

<u>Mutation</u>

Mutation operation was used in the proposed algorithm in order to generate a diversification effect. The inversion method was used to perform mutation.

The parameters of the proposed algorithm are population size, crossover fraction, mutation fraction, iteration bound.

Consequently, the aim of this study is to minimize the risk by using the proposed genetic algorithm.

5. The implementation and the obtained results

The most active 10 stocks traded in Istanbul Stock Exchange 30 Index (*ISE 30*) was chosen. Therefore, each chromosome has 10 genes. The used stocks are “Akbank”

(AKBNK), “Emlak Konut Gayrimenkul Yatırım Ortaklığı A.Ş.” (EKGYO), “Garanti Bankası” (GARAN), “Halk Bankası” (HALKB), “İpek Doğal Enerji Kaynakları Araştırma ve Üretim A.Ş.” (IPEKE), “İş Bankası C” (ISCTR), “Koza Anadolu Metal Madencilik İşletmeleri A.Ş.” (KOZAA), “Türkiye Petrol Rafinerileri A.Ş” (TUPRS), “Türkiye Vakıflar Bankası T.A.O.” (VAKBN), “Yapı Ve Kredi Bankası A.Ş.” (YAKBNK).

The series for each stocks contain weekly data between the dates 1 November 2011 and 31 October 2012. Totally 52 observations were taken as closure price for every week.

The parameters of the proposed algorithm used in the implementation are given in Table 1.

Table 1. **The parameters of the proposed genetic algorithm**

Parameter	***Values***
Population size	20
Crossover fraction	1
Mutation fraction	1
Iteration bound	100

As a result of the implementation, the obtained alternative solutions are presented in Table 2. In this table, the last two columns give corresponding expected return and risk values. After the proposed genetic algorithm is utilized, alternative solutions are obtained. Decision maker can choose the best solution based on his/her further personal criteria of assessment.

Table 2. **The obtained alternative solutions**

AKBNK	EKGYO	GARAN	HALKB	IPEKE	ISCTR	KOZAA	TUPRS	VAKBN	YKBNK	Expected Return	Risk
0.081	0.580	0.075	0.003	0.029	0.030	0.029	0.003	0.081	0.090	5	0.149
0.158	0.328	0.045	0.011	0.038	0.040	0.038	0.006	0.179	0.157	6	0.236
0.143	0.536	0.019	0.000	0.006	0.006	0.006	0.000	0.141	0.141	7	0.132
0.145	0.398	0.048	0.001	0.038	0.043	0.041	0.001	0.142	0.142	8	0.198
0.153	0.259	0.121	0.004	0.034	0.033	0.035	0.004	0.178	0.179	9	0.257
0.168	0.336	0.066	0.011	0.019	0.018	0.019	0.010	0.183	0.170	10	0.228
0.196	0.250	0.038	0.013	0.038	0.038	0.037	0.001	0.194	0.196	11	0.254
0.014	0.919	0.009	0.007	0.008	0.008	0.009	0.008	0.010	0.009	12	0.070
0.113	0.581	0.017	0.004	0.016	0.013	0.014	0.003	0.117	0.122	13	0.132
0.147	0.287	0.015	0.001	0.015	0.016	0.015	0.001	0.248	0.255	14	0.206
0.009	0.870	0.011	0.008	0.007	0.010	0.007	0.006	0.007	0.066	15	0.077
0.002	0.985	0.002	0.001	0.001	0.002	0.002	0.001	0.002	0.002	16	0.049
0.035	0.884	0.009	0.002	0.009	0.009	0.009	0.001	0.009	0.033	17	0.067

0.094	0.691	0.004	0.003	0.003	0.004	0.003	0.002	0.099	0.099	18	0.099
0.004	0.963	0.004	0.004	0.004	0.004	0.004	0.004	0.004	0.005	19	0.057
0.002	0.980	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.002	20	0.051

6. Conclusion

Portfolio optimization is a kind of knapsack problem that is a well-known optimization problem. Many researchers from different fields have been interested in solving this problem. The decision-maker's desire is to avoid the higher risks associated with larger monetary values when the decision maker has some investment opportunities. In this study, to provide support to the decision makers in the process of making a choice among different options, we propose an alternative solution approach which is based on genetic algorithms. The proposed genetic algorithm is introduced and it is applied to a real life problem. As a result of the implementation, sixteen different solutions with different risk levels and return values are obtained and presented.

References

Chang, T., J. et al. (2000) Heuristics for cardinality constrained portfolio optimisation. *Computers & Operations Research*, 27, 1271-1302.

Chang, T., J., Yang, S., C. and Chang, K., J. (2009) Portfolio optimization problems in different risk measures using genetic algorithm, *Expert Systems with Applications*, 36, 10529-10537.

Egrioglu, E., Aladag, C., H. and Yolcu, U. (2012) Comparison of feed forward and Elman neural networks forecasting ability: case study for IMKB, *Advances in Time Series Forecasting*, Bentham Science Publishers Ltd., 11-17.

Elton, J., E. and Gruber, J., M. (2011) *Investments and portfolio performance*, World Scientific, 382–383.

Gilli, M. and Këllezi, E. (2000) Heuristic approaches for portfolio optimization, *In*: *Sixth International Conference on Computing in Economics and Finance of the Society for Computational Economics*, Barcelona, July 6-8 2000.

J.,H., Holland (1975) *Adaptation in natural and artificial systems: An introductory analysis with applications to biology,control and artficial intelligence*, Michigan: University of Michigan Press.

Markowitz, H.(1959) *Portfolio selection: efficient diversification of ınvestments*, New York: Wiley.

Markowitz, H.(1987), *Analysis in portfolio choice and capital markets*, Oxford: Basil Blackwell.

Stancu, S. and Predescu, M.,O. (2010)/ Genetic algorithm for the portfolio selection problem on the Romanian capital market, *Proceedings of International Conference on Engineering an Meta-Engineering*, Orlando, USA, 57-60.

Zhang, B. and Hua, Z., S. (2012) Simple solution methods for separable mixed linear and quadratic knapsack problem, *Applied Mathematical Modelling*, 36, 3239-3250.