An Effective Decision-Based Genetic Algorithm Approach to Multiobjective Portfolio Optimization Problem

Chi-Ming Lin

Graduate School of Information, Production and Systems,
Waseda University, Kitakyushu, Japan
and
Center for General Education, College of Management,
Kainan University, Taoyuan, Taiwan

Mitsuo Gen

Graduate School of Information, Production and Systems,
Waseda University, Kitakyushu, Japan

Abstract

Multiobjective portfolio optimization problem is the portfolio process of the highest expected return among the various financial commodities of the capital market to meet the expected return objectives. And one of the most important and common management issues lies in determining the best portfolio out of a given set of investment proposals. As we know, modern portfolio theory provides a well-developed paradigm to form a portfolio with the highest expected return for a given level of risk tolerance. However, for making the profit via the limited available capital, allocating the money to construct a portfolio is a challenge to be dealt with. In the capital market, there are thousands of financial commodities. Depending on the characteristics of the commodity, the risk and return of the investment are dissimilar. And the risk and return should also be simultaneously considered in practice of stock market. Hence, portfolio optimization is a complex multiobjective problem of multistage decision-based. In this paper, the multistage decision-based genetic algorithm is proposed for the multiobjective portfolio optimization problem. On the basis of the illustrative example, we can show the effectiveness of the proposed algorithm is validated for solving this problem.

Keywords: Multiobjective Portfolio Optimization; Genetic Algorithm

1   Introduction

The multiobjective portfolio optimization problems have received increased interest from researchers with various backgrounds since early 1952. A modern portfolio theory provides a well-developed paradigm to form a portfolio with the highest expected return for a given level of risk tolerance. Markowitz (1952) [1], a creator of modern portfolio theory, originally formulated the fundamental theorem of mean-variance portfolio framework, which explains the trade-off between mean and variance each representing expected returns and risk of a portfolio, respectively. The mean-variance approach which proposed by Markowitz (1987) was to deal with the portfolio selection problem [2]. A decision-maker can determine the optimal investing ratio to each security based on the sequent return rate. The formulation of the mean-variance method can be described as follows:

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

$s.t. \sum_{i=1}^{n} \mu_i x_i \geq E, \sum_{i=1}^{n} x_i = 1, \, x_i \geq 0$ and $i = 1, 2, \ldots, n,$

where σ_{ij} denotes the covariance between the return of the ith security and the jth security, μ_i denotes the expected return rate of the ith security, E denotes the acceptable least rate of the expected return, x_i denotes the investment portion in the ith securities.

For a given specific return rate, one can derive the minimum investment risk by minimizing the variance of a portfolio; or for a given risk level which the investor can tolerate, one can derive the maximum returns by maximizing the expected returns of a portfolio. The main input data of the Markowitz mean-variance model are expected returns and variance of expected returns of these securities [5]. Although Markowitz's theory uses only mean and variance to describe the characteristics of return, his theory about the structures of a portfolio became a cornerstone of modern portfolio theory.

Due to the huge numbers of financial securities and the acceptable least rate of the expected return is difficult to estimate, the mean-variance model is not practical for applying. In the other hand, the investment activities are restricted the limited available capital. Some investors can not construct a profitable portfolio via a large amount of money. Hence, it is important to build up an efficient method to select the securities for constructing the optimal portfolio. As the result, we propose the multistage decision-based genetic algorithm approach for dealing with the multiobjective portfolio optimization problem. In our method, we tackle this problem through a two phase approach. Firstly, we select the short list of the securities by the past performance evaluation. Then, genetic algorithm is applied to decide the investment weight of the securities.

In the past few years, there has been a boom in applying genetic algorithms to solving the multiobjective optimization problem known as evolutionary multiobjective optimization or genetic multiobjective optimization. The basic feature of genetic algorithms is the multiple directional and global searches, in which a population of potential solutions is maintained from generation to generation. [3] The population-to-population approach is beneficial in the exploration of the securities optimal portfolio selecting solutions.

On the other hand, Genetic Algorithm (GA) and its application to various disciplines including optimization problems [1]. GA is also applied to a wide range of optimization, and can offer significant advantages in solution methodology and optimization performance. An useful feature of GA is to handle multiobjective function optimization [4].

The remainder of this paper is organized as follows. The solution approach including the genetic algorithm and the operators are discussed in Section 2. A numerical example is used to illustrate the proposed method and the results are shown in Section 3. And the conclusions are presented in Section 4.

2   Proposed Solution Approach

As discussed in Section 1, we propose a two phase method for solving the problem. Before introducing the approach, we have to discuss about the objectives of this problem. In this paper, we tackle the problem as a multiobjective optimization problem. The mathematical model, which is expanded form Markowitz mean-variance approach, is shown as follow:

Notation:
    Indices
        i, j: index of security, $i, j = 1, 2, \ldots, n$.
    Parameters
        μ_i: the expected of security, $i, j = 1, 2, \ldots, n$.
        σ_{ij}: the covariance between the returns of security i and security j,
            $i, j = 1, 2, \ldots, n$.
    Decision Variables
        w_i: investment weight of security i.
Mathematical Model:

$$\max \sum_{i=1}^{n} \mu_i w_i \tag{2}$$

$$\min \sum_{i=1}^{n} \sum_{j=1}^{n} \sigma_{ij} w_i w_j \tag{3}$$

s. t. $\sum_{i=1}^{n} w_i = 1$, $w_i \geq 0$ and $i = 1, 2, \ldots, n$.

The objective function (2) is to maximizing the total expected return. And the objective function (3) is to minimizing the total risk of the portfolio. In this model, we would like to track these two objective functions at the same time. And obtain a set of Pareto solutions after applying the multistage decision-based genetic algorithm approach. For solving this model, we assume the expected return μ_i is the past 12 months average return rate of security i and the covariance σ_{ij} is the covariance between μ_i and μ_j.

2.1   Candidate Securities Selection

Due to the financial market liberalization, there are more and more securities (or assets) in the capital market. In order to select the investment targets, the investors prefer securities which with outstanding performance. In our approach, we calculate the past 3-month performance. Then, choose the highest 20% for the candidate securities. And reorder the securities for the highest performance to the lowest. The 3-month performance of the security i, r_i^3, measures the relative change of the price of the security over the last three months (in percent) and is therefore a measure for the short term expected return. In particular,

$$r_i^3 = \frac{p_i^T - p_i^{T-3}}{p_i^{T-3}} \tag{4}$$

where p_i^T is the price of security i in period T and p_i^{T-3} is the price of security i in three months ago.

2.2   Genetic Algorithm Approach

Genetic Algorithm was developed initially by Holland (1975) form the 1960s. The usual form of GA was described by Goldberg (1989). GA is stochastic search technique based on the mechanism of natural selection and natural genetics. 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. Many research papers and dissertations have established the validity of GA approach in function optimization problems and application problems [4].

Genetic Algorithm, differing from conventional search techniques, 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. 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 [4].

2.2.1 Random Keys-base Encoding

The random keys-based encoding method is a direct approach, which encodes some information for constructing a set of weights in a chromosome. As we know, a gene in a chromosome is characterized by two factors: locus, (i.e., the position of the gene located within the structure of chromosome), and allele, (i.e., the value the gene takes). In this encoding method, the position of the gene is used to represent the ID number of the security and its value is used to represent the weight for constructing a portfolio. As proposed random keys-based encoding method, randomly generates the initial chromosome first. This encoding method is easily verified that any permutation of the encoding corresponds to the compositions of the portfolio, so that most existing genetic operators can easily be applied to the encoding. The pseudocode for order-based encoding as following.

More so than differs from other optimization techniques, GA provides a framework of using only objective function information for analyzing many problem types. Within this framework of optimization techniques can be employed to solve the non-smooth, non-continuous and non-differentiable functions which are actually existed in a practical optimization problem. [4]

2.2.2 Weight Generating

In the encoding procedure, the value of the gene in the chromosome is generated randomly. When we are generating the weight vector, we have to rescale the weight to satisfy $\sum_{i=1}^{n} w_i = 1$. As the result, we convert the weight into $w_i = \frac{v_i}{\sum_{i=1}^{N} v_i}$.

procedure 2: Weight Generating
Input: chromosome v,
            total number of securities N
output: weight vector W
begin
  $W \leftarrow \emptyset$,
  for $i = 1$ to N
  $$w_i = \frac{v_i}{\sum\limits_{i=1}^{n} v_i}$$
  $W \leftarrow W \cup w_i$;
  output weight set W;
end

Than an example of weight generating is shown in Fig. 1. We try to assign the weight for each security. The determination of these weights indicates the relative importance of the various objectives. The chromosome consists of vector of weight.

Security ID :	1	2	3	4	5	6	7	8
Chromosome $v_1(i)$:	0.52	0.35	0.44	0.68	0.56	0.23	0.11	0.19
Weight generating:	0.17	0.11	0.14	0.22	0.18	0.07	0.04	0.06

Figure 1: The example of weight generating

2.2.3 Crossover Operator

Crossover is the main genetic operator. It operates on two parents (chromosomes) at a time and generates offspring by combining both chromosomes' features. In weight selection problem, crossover plays the role of exchanging weights of the securities of two chosen parents in such a manner that the offspring produced by the crossover represents. Several crossover operators have been proposed for permutation representation, such as Partial-mapped crossover (PMX), Order crossover (OX), Position-based crossover (PX), heuristic crossover, and so on[3]. In this paper, we use a very simple crossover method, which called one-cut point crossover.
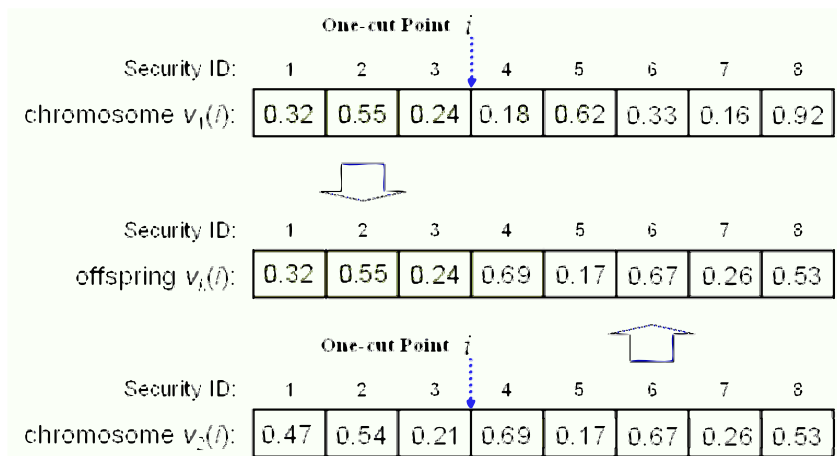
Figure 2: The Example of One-cut Point Crossover

2.2.4 Mutation Operator

Mutation is a background operator which produces spontaneous random changes in various chromosomes. A simple way to achieve mutation would be to alter one or more genes. In GA, mutation serves the crucial role of either replacing the genes lost from the population during the selection process, so they can be tried in a new context or providing the genes that were not present in the initial population. In this paper, it is relatively easy to produce some mutation operators for permutation representation.

Several mutation operators have been proposed for permutation representation, such as Swap mutation, Inversion mutation, and Insertion mutation, and so on (M. Gen and R. Cheng, 2000). In this case we use Insertion mutation method, and an example as following Fig. 3.
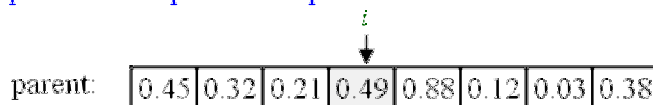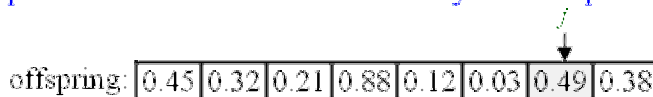

Figure 3: The Example of Insertion mutation

2.2.5 Evaluation

The evaluation function interprets the chromosome in terms of physical representation and evaluates its fitness based on traits of being desired in the solution. Evaluation function used for the GA is based on the total expected return and the risk of the portfolio. For the portfolio selection problem, we consider the total expected return and the risk. Therefore, the fitness function that involves computational efficiency and accuracy (of the fitness measurement) is defined as follows:

$$eval(v_k) = \frac{f(returen)}{f(risk)} = \frac{\sum_{i=1}^{n} \mu_i w_i}{\sum_{i=1}^{n} \sum_{j=1}^{n} \sigma_{ij} w_i w_j} \tag{5}$$

where $eval(v_k)$ represents the fitness value of the k-th chromosome.

2.2.6 Selection

Selection (reproduction) operator is intended to improve the average quality of the population by giving the high-quality chromosomes a better chance to get copied into the next generation. The principle behind genetic algorithms is essentially Darwinian natural selection. The selection directs GA search towards promising regions in the search space. We employ roulette wheel selection as a selection mechanism in this study. In roulette wheel selection mechanism, the individuals on each generation are selected for survival into the next generation according to a probability value proportional to the ratio of individual fitness over total population fitness; this means that on average the next generation will receive copies of an individual in proportion to the importance of its fitness value.

3 Numerical Example

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: $p_C = 0.70$;
Mutation probability: $p_M = 0.50$;

In this numerical example, 40 sample companies' data, which are collected form Taiwan's stock market, are used to demonstrate the proposed method.

Table 1: 3-months and 12-months return rates for 40 sample companies

ID	r_i^3	r_i^{12}	ID	r_i^3	r_i^{12}	ID	r_i^3	r_i^{12}	ID	r_i^3	r_i^{12}	ID	r_i^3	r_i^{12}
1	0.07	0.10	9	0.03	0.04	17	0.03	0.05	25	0.01	0.03	33	0.11	0.10
2	0.14	0.12	10	0.04	0.03	18	0.05	0.06	26	0.01	0.04	34	0.05	0.07
3	0.10	0.03	11	0.01	0.04	19	0.06	0.10	27	0.21	0.22	35	0.04	0.08
4	0.31	0.23	12	0.19	0.21	20	0.13	0.17	28	0.12	0.14	36	0.04	0.02
5	0.01	0.00	13	0.13	0.14	21	0.11	0.11	29	0.11	0.07	37	0.02	0.01
6	0.05	0.01	14	0.09	0.10	22	0.08	0.10	30	0.06	0.07	38	0.03	0.01
7	0.01	0.05	15	0.06	0.05	23	0.13	0.18	31	0.07	0.08	39	0.17	0.12
8	0.12	0.10	16	0.09	0.17	24	0.22	0.19	32	0.03	0.01	40	0.14	0.13

Next, we choose the candidate securities by the 3-months performance, and reorder the security ID number as follow:

Table 2: Reordering data sets

Reordering ID	ID	r_i^3	r_i^{12}
1	04	0.31	0.23
2	24	0.22	0.19
3	27	0.21	0.22
4	12	0.19	0.21
5	39	0.17	0.12
6	21	0.15	0.11
7	02	0.14	0.12
8	40	0.14	0.13

From Table 3, the variance-covariance matrix, we can get the data about the risk. Then, apply the proposed algorithm to construct the optimal portfolio. The best solution of the test problem is shown as Table 4.

4 Conclusions

Mean-variance is widely used in the finance area to deal with the portfolio selection problem. The purpose of the mean-variance approach is to determine the period optimal investing rate to each security based on the sequent return rate.

A genetic algorithm is designed to solve the corresponding optimization problems because these nonconcave maximization problems are with a particular structure and cannot be efficiently solved by the existing traditional optimization methods. In this paper, the multiobjective optimal portfolio selection problem is considered. We construct a two phases approach, and apply

Table 3: The covariance matrix

Covariance matrix	1	2	3	4	5	6	7	8
1	1.231	0.028	1.558	1.506	1.440	1.312	1.445	1.024
2	0.028	1.114	-0.074	-0.024	0.034	0.013	-0.005	-0.022
3	1.558	-0.074	2.081	1.847	1.525	1.143	1.096	1.002
4	1.506	-0.024	1.847	1.234	1.468	1.679	1.548	1.332
5	1.440	0.034	1.525	1.468	1.113	1.022	1.008	1.321
6	1.312	0.013	1.143	1.679	1.022	1.020	1.431	1.030
7	1.445	-0.005	1.096	1.548	1.008	1.431	1.103	1.111
8	1.024	-0.022	1.002	1.332	1.321	1.030	1.111	1.301

Table 4: The optimal solution

The weight vector of optimal portfolio	Total expected return rate	Risk
{0.34, 0.22, 0.13, 0.09, 0.10, 0.02, 0.04, 0.06}	0.1943	0.898

it for the test problem. As the experiment result, the proposed method is valid for the portfolio optimization problem.

References

[1] H. Markowitz, Portfolio Selection, *Journal of Finance,* **7** (1952), 77 - 91.

[2] H. Markowitz, *Mean-Variance Analysis in Portfolio Choice and Capital Market,* Basil Blackwell, New York, 1987.

[3] M. Gen and R. Cheng, *Genetic Algorithms and Engineering Design,* John Wiley, New York, 1997.

[4] M. Gen and R. Cheng, *Genetic Algorithms and Engineering Optimization,* John Wiley, New York, 2000.

[5] M. Ehrgott, K. Klamroth and C. Schwehm, An MCDM approach to portfolio optimization, *European Journal of Operational Research,* **155** (2004), 752 - 770.