

O.R. Applications

Genetic algorithms for portfolio selection problems with minimum transaction lots

Chang-Chun Lin ^{a,*}, Yi-Ting Liu ^b

^a Department of Information Management, Chang Gung University, 259 Wen-Hwa 1st Road, Kwei-Shan, Tao-Yuan 333, Taiwan, ROC

^b Institute of Information Management, National Chiao Tung University, 1001 Ta Hsueh Road, Hsinchu 300, Taiwan, ROC

Received 12 March 2006; accepted 11 December 2006

Available online 11 January 2007

Abstract

Conventionally, portfolio selection problems are solved with quadratic or linear programming models. However, the solutions obtained by these methods are in real numbers and difficult to implement because each asset usually has its minimum transaction lot. Methods considering minimum transaction lots were developed based on some linear portfolio optimization models. However, no study has ever investigated the minimum transaction lot problem in portfolio optimization based on Markowitz' model, which is probably the most well-known and widely used. Based on Markowitz' model, this study presents three possible models for portfolio selection problems with minimum transaction lots, and devises corresponding genetic algorithms to obtain the solutions. The results of the empirical study show that the portfolios obtained using the proposed algorithms are very close to the efficient frontier, indicating that the proposed method can obtain near optimal and also practically feasible solutions to the portfolio selection problem in an acceptable short time. One model that is based on a fuzzy multi-objective decision-making approach is highly recommended because of its adaptability and simplicity.

© 2007 Published by Elsevier B.V.

Keywords: Portfolio selection; Markowitz model; Minimum transaction lots; Genetic algorithm; Fuzzy multi-objective decision making

1. Introduction

Markowitz' model [1] uses the mean and variance of historical returns to measure the expected return and risk of a portfolio. Conventionally, such portfolio selection problems are solved with quadratic or linear programming models under the assumption that the asset weights in the portfolio are real num-

bers, which are difficult to implement. Specifically, each asset has its minimum transaction lot, while the solutions involve only real-number asset weights rather than asset trading units. For example, stocks might be traded at the unit one share, and mutual funds have their individual minimum trading amounts. Thus, the solution obtained by Markowitz' model must be integers to be applicable in practice.

Other than Markowitz' model, Speranza [2], Mansini and Speranza [3,4], and Kellerer et al. [5] proposed their respective portfolio selection models

* Corresponding author. Tel.: +886 32118800; fax: +886 32118020.

E-mail address: chanclin@ms29.hinet.net (C.-C. Lin).

based on Konno and Yamazaki's mean absolute deviation (MAD) model [6]. Speranza [2] proposed a mixed integer program considering realistic characteristics in portfolio selection, such as minimum transaction lots and the maximum number of securities, and suggested a simple two-phase heuristic algorithm to solve the proposed integer program. Mansini and Speranza [4] showed that the portfolio selection problem with minimum transaction lots is an NP-complete problem and proposed three heuristic algorithms to solve the problem. Based on the MAD model, Konno and Wijayanayake [7] proposed an exact algorithm for portfolio optimization problems under concave transaction costs and minimum transaction lots. However, minimum transaction lots were not the major concern in their study. Later, Mansini and Speranza [8] derived a mean-safety model with side constraints from the MAD model, and proposed an exact algorithm to solve for portfolios under the consideration of transaction costs and minimum transaction lots. However, Markowitz' model is still the most widespread portfolio selection model. Solving the portfolio selection problem based on Markowitz' model and, simultaneously, considering minimum transaction lots are of practical significance. However, it appears that no methods in the past solving the portfolio selection problem with minimum transaction lots were based on Markowitz' model.

Apart from considering minimum transaction lots, Markowitz' model is intrinsically a multi-objective decision-making (MODM) problem whose decision criteria conflict with each other. The return is required to be maximized and the risk minimized. However, the risk is often high when the return is maximized, and the return is often low when the risk is minimized. Researchers have proposed different approaches, such as goal programming [10] and multiple objective programming [11], to solve multi-objective portfolio selection problems. Lee and Lerro [10] pioneered goal programming in portfolio selection, but their method is not based on Markowitz' model. Arenas Parra et al. [12] proposed a fuzzy goal programming approach to solve a portfolio selection problem, using a multi-index model to estimate the return and risk of portfolios and treating fuzzy goals as fuzzy numbers. However, their method ignores the existence of minimum transaction lots. Apart from goal programming, fuzzy programming can also be used to solve MODM problems [13], but conventional fuzzy programming methods cannot incorpo-

rate objective weights that are usually important for the decision maker (DM) to express his/her preference regarding return and risk. Lin [14] recently proposed a weighted max–min model to incorporate objective weights with fuzzy programming. Promisingly, this method can be applied to solving the portfolio selection problem. Using this method, the DM obtains portfolios through setting appropriate objective weights.

The portfolio selection problem with minimum transaction lots is a combinatorial optimization problem whose feasible region is not continuous. Studies have shown that genetic algorithm (GA) is a promising approach to combinatorial optimization problems. Shoaf and Foster [9] applied GA to Markowitz' portfolio selection problem and found that the time complexity of GA approximates $O(n \log n)$ and is better than that of quadratic programming. However, their approach does not consider the constraints of minimum transaction lots. The solution obtained will be much more realistic if the constraints of minimum transaction lots are considered when solving with GAs. Therefore, this study presents and compares three possible decision models with their corresponding GAs to solve the portfolio selection problem with minimum transaction lots, an NP-complete MODM problem. The proposed GAs are highly efficient and effective in providing near optimal solutions within a few minutes.

The rest of this paper is organized as follows. Based on Markowitz' model, Section 2 derives two models for portfolio selection with minimum transaction lots. Section 3 presents the fuzzy MODM model of Lin [14], and then an integrated model for portfolio selection based on fuzzy MODM and considering minimum transaction lots is proposed. The GAs for solving the proposed models are stated in Section 4. Section 5 presents the experiments with the proposed GAs on the mutual funds in Taiwan and discusses the results. Finally, Section 6 concludes the paper and suggests possible future research.

2. Portfolio selection and minimum transaction lots

Originally an MODM problem, portfolio selection attempts to maximize the rate of return and minimize the portfolio risk simultaneously. In practice, the MODM problem is often degenerated to a single objective one by introducing a preference structure to compromise between the objectives, or simply optimizing one of the two objectives while

bounding the remaining one. The latter approach leads to the well-known Markowitz’ model [1]. Markowitz’ model assumes that rational investors aim to maximize return under a certain risk level or minimize risk above a certain return level. For convenience, the decision maker usually fixes the expected rate of return and then minimizes the portfolio risk under this return constraint. Consequently, Markowitz’ model can be expressed as a quadratic programming problem as follows:

Model 1

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

$$\text{subject to} \quad r_p = \sum_{i=1}^n w_i r_i = r, \quad (2)$$

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

where n is the number of assets; r_p is the expected rate of return; r is the required rate of return; r_i is the expected rate of return of asset i ; σ_{ij} is the return covariance between assets i and j , and σ_p^2 is the return variance of the portfolio. w_i represents the weight of budget invested in asset i . As in the original Markowitz’ model, $w_i \geq 0$, that is, no short selling is allowed.

However, in practice, each asset has its minimum transaction lot that should be taken into consideration in finding minimum-risk portfolios. Solutions to portfolio selection problems thus must be integers because real-number solutions might be difficult to implement. As a consequence, the portfolio selection models need to be modified to consider the minimum transaction lots. A modified Markowitz’ model can be formulated as follows to consider the minimum transaction lots.

Model 2

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

$$\text{subject to} \quad e_p = \sum_{i=1}^n x_i c_i r_i \geq br, \quad (4)$$

$$\sum_{i=1}^n x_i c_i \leq b, \quad (5)$$

$$w_i = \frac{c_i x_i}{\sum_{i=1}^n c_i x_i}, \quad i = 1, \dots, n, \quad (6)$$

$$x_i \in Z,$$

where b represents the budget and e_p the expected portfolio return. c_i denotes the unit price of asset i , and x_i represents the units invested in asset i . Since the same rate of return can be achieved with different budgets, Eq. (4) uses portfolio return instead of rate of return to maximally utilize the budget. More return is earned as more capital is invested. Nevertheless, constraint (5) requires that the total investment be below the budget. Similar to (3), constraint (6) requires that the sum of weights be unity. Notably, this constraint causes the model to be nonlinear.

Similar to Model 1, Model 2 minimizes the investment risk with respect to a given rate of return r except that the optimal solution must be in integers. However, the modification might make a portfolio with a rate of return equals r unobtainable, namely, the rate of return might not exactly equal r . Owing to constraints (4), the rate of return can only exceed r . In a portfolio selection problem whose solution is in real numbers, the desired rate of return along with its corresponding risk represents a target portfolio to be achieved. The target portfolio can always be achieved in a real-number solution space. However, in an integer solution space, the target portfolio might be unreachable, and the optimal solution conforms to constraint (4) can be far from the target. Specifically, the risk of the optimal solution to Model 2 might be too high to be acceptable if its rate of return is to exceed the desired r . Conversely, a portfolio with a rate of return slightly less than the desired r might be more appealing if the portfolio risk can accordingly be lowered significantly. This alternative leads to another decision model that minimizes the distance between the obtained and the target portfolios. The model is as follows:

Model 3

$$\min \quad d^2 = \left(\frac{\sigma_p - \sigma}{\sigma^* - \sigma} \right)^2 + \left(\frac{e_p - br}{br^* - br^-} \right)^2, \quad (7)$$

$$\text{subject to} \quad \sigma_p^2 = \sum_{i=1}^n \sum_{j=1}^n w_i w_j \sigma_{ij}, \quad (1)$$

$$e_p = \sum_{i=1}^n x_i c_i r_i, \quad (8)$$

$$\sum_{i=1}^n x_i c_i \leq b, \quad (5)$$

$$w_i = \frac{c_i x_i}{\sum_{i=1}^n c_i x_i}, \quad i = 1, \dots, n, \quad (6)$$

$$x_i \in Z,$$

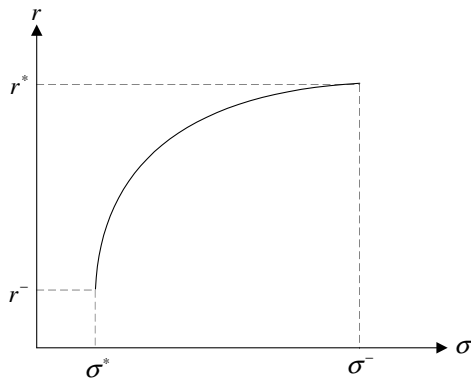


Fig. 1. Portfolio efficient frontier.

where σ is the minimum risk corresponding to the desired r . Notably, the risk and return are normalized to eliminate the influence of scales. As shown in Fig. 1, r^* and σ^* represent the maximal rate of return and the minimal risk that can be achieved, and σ^- and r^- are their corresponding risk and rate of return. Therefore, (σ^*, r^-) and (σ^-, r^*) represent the lower and upper ends of the efficient frontier.

Obviously, this model is also a nonlinear integer programming problem that is difficult and time-consuming to solve. Fortunately, genetic algorithms are helpful in solving difficult nonlinear and/or combinatorial optimization problems. That is why this study uses genetic algorithms to solve the portfolio selection problem with minimum transaction lots. Nevertheless, despite seemingly workable, the models still have some potential problems. The same rate of return might involve different levels of risk in different markets or in different time periods. Determining an appropriate value of r to reflect the investor risk preference is a tricky business. Furthermore, the desired σ for Model 3 must be obtained in advance by solving Markowitz' model (Model 1). Obtaining the target portfolio might be time-consuming when the number of assets is large. The next section presents an alternative approach that might be more applicable to solving the portfolio selection problem, which is a fuzzy MODM approach requiring no predetermined r and σ , and might be promising when applied to different scenarios.

3. Fuzzy multi-objective portfolio selection

Conventional MODM approaches involve goal programming, compromise programming and fuzzy programming. Goal programming minimizes the weighted sum of absolute deviations to the objective

goal values subjectively given by the DM. Model 3 can be considered as a goal programming approach. The major problem with goal programming lies in the incommensurability between objectives of different measures. Different metrics of objectives lead to incommensurable goal values and deviations from them. Summing up these incommensurable deviations is simply unreasonable. Compromise programming minimizes the sum of weighted distances to the ideal solution, based on some distance metric. Usually, compromise programming uses Minkowsky distance L_p as the distance metric, and thus, the solution obtained depends on the parameter p . Although the incommensurability problem in compromise programming can be resolved by normalization, the choice of an appropriate distance parameter p becomes a tricky problem to the DM. Fuzzy programming has been widely and successfully applied to various MODM problems, such as those in the studies of Chang and Chen [15], Abd El-Wahed and Abo-Sinna [16], and Rasmy et al. [17]. Fuzzy programming helps to resolve the incommensurability problem between objectives, and thus provides an alternative for solving MODM problems.

Since Zadeh [18] proposed fuzzy set theory in 1965, related theories and methodologies have been widely applied to many fields, including operations research, decision science, engineering and artificial intelligence. Fuzzy programming arose when Zimmermann [13] first applied the max–min operator of Bellman and Zadeh [19] to MODM problems. The DM defined a membership function for each objective to represent the achieved level of that objective regarding different objective values. Fuzzy programming thus maximizes the minimal achieved level among the objectives. Narasimhan [20] incorporated fuzzy sets with goal programming before many related approaches for such problem, fuzzy goal programming, were proposed. Fuzzy programming can be regarded as a special case of fuzzy goal programming where the goals to achieve are precisely the optimal values for the objectives.

A fundamental fuzzy goal programming problem with m fuzzy goals can be formulated as follows:

Model 4

$$\text{Find } \mathbf{x}, \quad (9)$$

$$\text{to satisfy } f_j(\mathbf{x}) \widetilde{\geq} g_j, \quad j = 1, 2, \dots, m, \quad (10)$$

$$\text{subject to } \mathbf{B}\mathbf{x} \leq \mathbf{b}, \quad (11)$$

$$\mathbf{x} \geq 0,$$

where \mathbf{x} is the solution in vector form with components x_1, x_2, \dots, x_n ; $f_j(\mathbf{x})$ means the j th objective, and $\mathbf{B}\mathbf{x} \leq \mathbf{b}$ are the system constraints in vector notation. The notation \geq means that the relation \geq is fuzzy such that the goal g_j can be partially achieved. The achievement of g_j is measured by a membership function defined for that goal. Any fuzzy goal programming problem can be expressed with Model 4 without any loss of generality because every \cong constraint can be replaced with a \leq constraint and a \geq constraint, and every \leq constraint can be converted to an equivalent \geq constraint.

Since all objectives might not be fully achieved simultaneously, the DM may define a lower tolerance limit and a corresponding membership function for each objective to determine the achieved level of that objective. A membership function $\mu_j(f_j)$ for the j th goal can be defined as

$$\mu_j(f_j) = \begin{cases} 1 & \text{if } g_j \leq f_j(\mathbf{x}), \\ \frac{f_j(\mathbf{x}) - l_j}{g_j - l_j} & \text{if } l_j \leq f_j(\mathbf{x}) < g_j, \\ 0 & \text{if } f_j(\mathbf{x}) < l_j, \end{cases} \quad (12)$$

where g_j and l_j are the specified goal value and the lower tolerance limit for the j th objective.

After Zimmermann [13], several alternative methods [21–23] have been proposed to solve fuzzy MODM problems. However, no model seems to consider objective weights. Lin [14] recently proposed a weighted max–min model for fuzzy MODM problems. The weighted max–min model can be expressed as follows:

Model 5

$$\max \quad \lambda, \quad (13)$$

$$\text{subject to } \varpi_j \lambda \leq \frac{f_j(\mathbf{x}) - l_j}{g_j - l_j}, \quad j = 1, 2, \dots, m, \quad (14)$$

$$\mathbf{B}\mathbf{x} \leq \mathbf{b}, \quad (11)$$

$$\mathbf{x} \geq 0.$$

where λ represents the minimal achieved level among the objectives; ϖ_j denotes the weight of the j th objective.

Lin’s model can readily be applied to the portfolio selection problem. Let return and risk be the first and second objectives. Then let br^* and σ^* be their goal values, and br^- and σ^- their tolerance limits. The membership functions for return and risk can be defined as follows, respectively.

$$\mu_r(e_p) = \begin{cases} 1 & \text{if } br^* \leq e_p, \\ \frac{e_p - br^-}{br^* - br^-} & \text{if } br^- \leq e_p < br^*, \\ 0 & \text{if } e_p < br^-, \end{cases} \quad (15)$$

$$\mu_\sigma(\sigma_p) = \begin{cases} 1 & \text{if } \sigma_p \leq \sigma^*, \\ \frac{\sigma^- - \sigma_p}{\sigma^- - \sigma^*} & \text{if } \sigma^* \leq \sigma_p < \sigma^-, \\ 0 & \text{if } \sigma^- < \sigma_p. \end{cases} \quad (16)$$

Let ϖ_r and ϖ_σ denote the weights of return and risk, respectively. The fuzzy programming model for the portfolio selection problem can be formulated as Model 6.

Model 6

$$\max \quad \lambda, \quad (13)$$

$$\text{subject to } \varpi_r \lambda \leq \frac{e_p - br^-}{br^* - br^-}, \quad (17)$$

$$\varpi_\sigma \lambda \leq \frac{\sigma^- - \sigma_p}{\sigma^- - \sigma^*}, \quad (18)$$

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

$$e_p = \sum_{i=1}^n c_i r_i x_i, \quad (8)$$

$$\sum_{i=1}^n x_i c_i \leq b, \quad (5)$$

$$w_i = \frac{c_i x_i}{\sum_{i=1}^n c_i x_i}, \quad i = 1, \dots, n. \quad (6)$$

$$x_i \in Z.$$

With Model 6, the DM expresses his/her preferences for the return and risk through assigning the objectives weights. DMs with higher risk aversion can give the risk higher weights to raise its achieved level. The portfolio risk can be lowered through sacrificing the achieved level of return. By contrast, DMs with lower risk aversion give the return higher weights to raise its achieved level. Assume, without losing generality, that $\varpi_r + \varpi_\sigma = 1$. $\varpi_r = 1$ indicates that the risk-neutral DM considers only return, and the left-hand side of constraint (18) equals zero, representing that any risk is acceptable. By contrast, $\varpi_\sigma = 1$ indicates that the DM considers only risk and thus that any return is acceptable. The DM can easily obtain a satisfied efficient portfolio through assigning proper weights to the objectives. Model 6 still represents a time-consuming nonlinear combinatorial optimization problem such that a genetic algorithm is necessary for solution.

4. Genetic algorithms

Many optimization problems are very complex and difficult to solve by conventional methods. Therefore, evolutionary algorithms simulating natural processes were developed to solve them [24]. Among them, GA has been successfully in solving many optimization problems. Originally proposed by Holland [25], GA is a stochastic searching technique based on the mechanism of genetics and natural selection [26]. The models for portfolio selection with minimum transaction lots are combinatorial optimization problems. Many studies have showed that GAs can efficiently find near optimal or even the optimal solutions for many combinatorial optimization problems. Most GAs, including those in this study, use roulette wheel selection, one-point crossover and one-point mutation [25]. The roulette wheel approach belongs to the fitness-proportional selection and can select a new population with respect to the probability distribution based on fitness values. The GAs in this study employs a different evolution mechanism from conventional ones. In the original GA [25], all parents are replaced by their offspring to form a new generation, which is called generational replacement. Offspring may be less fit than their parents because GAs are blind, causing some fitter chromosomes to be lost from the evolutionary process. Therefore, several replacement processes have been examined to resolve this problem. Holland [25] suggested to replace an arbitrarily selected chromosome with a new-born offspring. DeJong [27] proposed a crowd-

ing strategy that selects the parent that is most closely resembles the new-born offspring to die when the offspring is born. Gen and Cheng [28] proposed that both parents and their immediate offspring are all candidates for the new generation to preserve the best chromosome. The GAs in this study follows a similar strategy to prevent the populations from degenerating. The population evolves continually without being replaced by another population. Let $P(t)$ be the population at iteration t . Fig. 2 describes the basic GA.

Although the GA represents a possible way of solving the models, some problems remain in its implementation. The main problem in applying a GA to constrained optimization problems is how to deal with the constraints. Constraints can be dealt with strategies such as reject, repairing and penalty strategies, and the strategy of modifying genetic operators [28]. The reject strategy excludes infeasible solutions immediately on generation, resulting in an efficient GA. The repairing strategy transforms an infeasible solution into a feasible one through a repairing process. The difficulty in designing a repairing process to comply with the problem weakens the repairing strategy. The penalty strategy uses a penalty function to penalize all infeasible solutions, hoping that infeasible solutions might evolve toward feasible. Finally, the strategy of modifying genetic operators aims to devise problem-specific representations and specialized genetic operators to maintain feasibility. Comparatively, the strategies of penalty and modifying genetic operators appear more suitable for this study. Therefore,

```

begin
   $t \leftarrow 0$ ;
  initialize  $P(t)$ ;
  evaluate  $P(t)$ ;
  find the best and worst chromosomes of  $P(t)$ ;
  while  $t <$  a predetermined iteration number do
    select two parents from  $P(t)$ ;
    generate two offspring by crossover and mutation;
    evaluate the offspring;
    if the offspring is fitter than the worst chromosome of  $P(t)$  then
      randomly select a parent except the best one and replace it;
    find the best and worst chromosomes of  $P(t)$ ;
     $t \leftarrow t + 1$ ;
  end
end

```

Fig. 2. Proposed genetic algorithm.

the strategies of penalty and modifying genetic operators were used in this study to deal with constraints (4) and (5), respectively.

4.1. Encoding and decoding

The way to encode a solution into a chromosome is a key issue in using GAs. Although the solutions to the proposed models are integers, this study opted to use real numbers, for the sake of operational simplicity, to encode the weight w_i instead of directly encoding the trading units x_i with an integer. The genes of a chromosome are real numbers between 0 and 1 to represent the weights invested in the assets. However, the summation of these weights might not be 1 in the initialization stage or after genetic operations and thus violate constraint (6). To overcome this problem, the weights are normalized as follows:

$$w'_i = \frac{w_i}{\sum_{i=1}^n w_i}, \quad (19)$$

where w'_i represents the new weight invested in asset i after normalization. Even though, the normalized weights are still inapplicable, as stated previously. Multiplying w_i by the budget b gives the actual amount that could be invested in asset i . However, there are usually residuals left. The actual units invested in asset i is then obtained by

$$x_i = \left\lfloor \frac{bw'_i}{c_i} \right\rfloor, \quad i = 1, \dots, n. \quad (20)$$

The residuals make the total investment below the budget b and thus conformant with constraint (5). However, the actual investment weights are prone to change. Eq. (6) is used to compute the actual weights before computing the return variance of the portfolio. The w_i of the chromosomes in the initial population is randomly generated before being normalized using Eq. (19). Two of the three effective constraints in the models, constraints (5) and (6), can be satisfied with the above weight normalizing and budget controlling strategies. However, constraint (4) in Model 2 might be easier to handle using a penalty strategy, which will be stated in the next subsection.

4.2. Fitness functions

Three GAs were implemented to verify the effectiveness of Models 2 and 3 and 6. The primary differences between these models lie in their objective

functions, leading to different fitness functions in their respective GAs. With the encoding and decoding scheme, the constraints in Models 2 and 3 are resolved, leaving constraint (4) in Model 2 to be dealt with a penalty strategy. The penalty function that impels the solutions to satisfy constraint (4) is formulated as follows to make the actual return exceed the expected return.

$$p(\mathbf{x}) = \begin{cases} br - \sum_{i=1}^n x_i c_i r_i & \text{if } br > \sum_{i=1}^n x_i c_i r_i, \\ 0 & \text{otherwise.} \end{cases} \quad (21)$$

Incorporating the objective function (1) and the penalty function (21), the fitness function for Model 2 can be defined as,

$$\text{fitness} = \exp(-k(\sigma_p^2 + Mp(\mathbf{x}))), \quad (22)$$

where k is a positive constant and M a large positive number. The negative exponent transforms the minimization problem into an equivalent maximization problem for the GA to solve. The exponential function with constant k confines the range of the fitness and thus alleviates the selection pressure of chromosomes with higher fitness, to prevent the GA from premature convergence. The large positive number M forces the solution to meet constraint (4) before minimizing the portfolio risk.

Model 3 is also a minimization problem such that the fitness function for this problem is defined similarly as,

$$\text{fitness} = \exp(-kd^2). \quad (23)$$

The fitness function for Model 6, a maximization problem, is simply its objective function λ , and thus

$$\text{fitness} = \min \left\{ \frac{e_p - br^-}{\varpi_r(br^* - br^-)}, \frac{\sigma^- - \sigma_p}{\varpi_\sigma(\sigma^- - \sigma^*)} \right\}. \quad (24)$$

5. Empirical study

This study used Taiwanese mutual fund data from the year 1997 to 2000 to test the proposed models and GAs. The monthly rates of return were used to determine the mean rates of return of the mutual funds and the return covariances between the assets in each year. The monthly rates of return were used instead of weekly ones because monthly horizon is more likely to be the real investment horizon than weekly horizon when investing in mutual funds. Studies [29,30] have shown that portfolio performance indexes such as Sharpe's measure do

depend on sampling horizon and thus its estimated risk measure. However, while a simple experiment has been conducted to find out that the observed risks of weekly data tend to be significantly larger than those of monthly data, whether sampling horizon affects the obtained portfolios remains uninvestigated. Therefore, this study has chosen to adopt monthly data.

Also, the r^* , r^- , σ^* and σ^- for each year's data were obtained with Model 1 and LINGO. Table 1 shows the characteristic values of the terminal portfolios on each year's efficient frontier. To measure and compare the effectiveness of the proposed models, a number of target portfolios spread on the efficient frontiers are selected for the GAs to approach. For Models 2 and 3, let the desired rate of return r be $\alpha r^- + (1 - \alpha)r^*$. Eleven values of α ranging from 0 to 1 and differenced by 0.1 were used. Model 2 was then used to obtain the optimal σ with respect to these target rates of return. Table 2 shows the 11 target portfolios for each year. However, the desired rate of return disappears in Model 6. The solution to this model depends upon the objective weights $(\varpi_r, \varpi_\sigma)$. Since each target portfolio corresponds to a specific set of $(\varpi_r, \varpi_\sigma)$, the corresponding objective weights are derived inversely from the 11 target portfolios to be used in Model 6 and its corresponding GA. From Eqs. (17) and (18), for a target portfolio on the efficient frontier there would be

$$\frac{\varpi_r}{\varpi_\sigma} = \frac{\mu_r}{\mu_\sigma} \tag{25}$$

Let $\varpi_\sigma = 1 - \varpi_r$. Eq. (25) leads to,

$$\frac{\varpi_r}{(1 - \varpi_r)} = \frac{br_p - br^-}{\frac{br^* - br^-}{\sigma^- - \sigma_p}} \tag{26}$$

Thus, the ϖ_r for a target portfolio with (σ, r) can be obtained by

$$\varpi_r = \frac{\frac{r - r^-}{r_p - r^-} - \frac{r^* - r^-}{\sigma^- - \sigma}}{\frac{r^* - r^-}{r_p - r^-} + \frac{\sigma^- - \sigma}{\sigma^- - \sigma^*}} \tag{27}$$

Table 1
Characteristics of efficient frontiers

Year	Number of funds	r^* (%)	r^- (%)	σ^* (%)	σ^- (%)
1997	129	6.663	0.480	0.007	15.296
1998	160	3.674	0.480	0.014	7.362
1999	197	9.158	0.388	0.014	8.925
2000	204	0.935	0.422	0.007	1.557

Table 2
Characteristics of target portfolios

Year	ϖ_r	r (%)	σ (%)
1997	0.000	0.480	0.007
	0.097	1.098	1.076
	0.189	1.716	2.210
	0.277	2.335	3.344
	0.364	2.953	4.600
	0.450	3.571	5.955
	0.536	4.189	7.342
	0.620	4.808	8.745
	0.704	5.426	10.159
	0.787	6.044	11.582
	1.000	6.663	15.296
1998	0.000	0.480	0.014
	0.097	0.799	0.538
	0.192	1.119	1.171
	0.284	1.438	1.806
	0.374	1.758	2.442
	0.462	2.077	3.077
	0.547	2.396	3.713
	0.631	2.716	4.348
	0.712	3.035	4.984
	0.795	3.355	5.657
	1.000	3.674	7.362
1999	0.000	0.388	0.014
	0.096	1.265	0.581
	0.188	2.142	1.232
	0.275	3.019	1.883
	0.358	3.896	2.534
	0.437	4.773	3.185
	0.512	5.650	3.836
	0.584	6.527	4.487
	0.656	7.404	5.181
	0.788	8.281	6.767
	1.000	9.158	8.925
2000	0.000	0.422	0.007
	0.092	0.474	0.024
	0.169	0.525	0.036
	0.251	0.576	0.168
	0.337	0.627	0.337
	0.430	0.679	0.530
	0.530	0.730	0.732
	0.636	0.781	0.936
	0.749	0.832	1.143
	0.871	0.884	1.350
	1.000	0.935	1.557

Table 2 also shows the corresponding objective weights of the target portfolios.

The parameter setting of the proposed GAs is as follows. For each year's data, the number of population and the chromosome length both equal the fund number. The crossover and mutation rates were set to 1 and 0.05, respectively. The number of iterations was 5000. The values of k and the M were set to 0.005 and 100, respectively. The budget

Table 3
Comparisons between models

Year	Model 4			Model 5			Model 9		
	r_p (%)	σ_p (%)	Deviation (%)	r_p (%)	σ_p (%)	Deviation (%)	r_p (%)	σ_p (%)	Deviation (%)
1997	0.494	0.012	0.015	0.494	0.013	0.015	0.486	0.012	0.008
	1.155	1.250	0.182	1.117	1.177	0.102	1.159	1.249	0.183
	1.789	2.453	0.253	1.729	2.330	0.121	1.769	2.400	0.197
	2.404	3.561	0.228	2.357	3.457	0.115	2.378	3.523	0.184
	3.073	5.026	0.442	2.979	4.766	0.168	2.981	4.908	0.310
	3.703	6.476	0.537	3.555	6.149	0.195	3.611	6.235	0.282
	4.299	7.821	0.491	4.191	7.543	0.201	4.168	7.599	0.258
	4.873	9.167	0.427	4.804	8.946	0.200	4.755	8.879	0.144
	5.514	10.509	0.361	5.420	10.318	0.159	5.403	10.239	0.083
	6.091	11.750	0.175	6.049	11.680	0.098	6.007	11.613	0.049
6.663	15.296	0.000	6.650	15.280	0.021	6.663	15.296	0.000	
Average			0.283			0.127			0.154
1998	0.508	0.048	0.044	0.499	0.046	0.037	0.503	0.043	0.037
	0.853	0.701	0.172	0.806	0.623	0.086	0.854	0.710	0.181
	1.196	1.401	0.242	1.138	1.272	0.102	1.175	1.336	0.174
	1.538	2.071	0.283	1.461	1.922	0.118	1.512	2.027	0.233
	1.865	2.707	0.287	1.767	2.552	0.111	1.811	2.643	0.208
	2.190	3.418	0.359	2.117	3.259	0.186	2.106	3.271	0.196
	2.504	4.013	0.319	2.429	3.853	0.144	2.444	3.893	0.187
	2.845	4.683	0.359	2.752	4.521	0.176	2.696	4.451	0.105
	3.154	5.339	0.374	3.064	5.086	0.106	2.995	5.075	0.099
	3.457	6.274	0.626	3.357	5.767	0.110	3.240	5.524	0.176
3.674	7.362	0.000	3.674	7.362	0.000	3.674	7.362	0.000	
Average			0.279			0.107			0.145
1999	0.415	0.016	0.027	0.388	0.016	0.002	0.415	0.016	0.027
	1.454	0.866	0.342	1.253	0.708	0.127	1.406	0.840	0.294
	2.394	1.620	0.463	2.183	1.407	0.179	2.328	1.542	0.361
	3.357	2.375	0.597	2.987	2.071	0.190	3.200	2.259	0.417
	4.320	3.234	0.818	3.832	2.734	0.209	3.784	2.978	0.458
	5.138	4.053	0.941	4.761	3.554	0.369	4.760	3.626	0.441
	5.979	4.606	0.838	5.479	4.229	0.428	5.458	4.267	0.471
	7.031	5.440	1.078	6.372	4.900	0.441	6.281	4.791	0.391
	7.855	6.038	0.968	7.339	5.647	0.470	7.015	5.393	0.443
	8.980	8.659	2.017	7.955	6.737	0.327	7.929	6.056	0.793
9.158	8.925	0.000	9.158	8.925	0.000	9.158	8.925	0.000	
Average			0.735			0.249			0.372
2000	0.434	0.019	0.016	0.436	0.020	0.019	0.400	0.013	0.023
	0.486	0.051	0.029	0.485	0.050	0.028	0.487	0.051	0.031
	0.538	0.168	0.133	0.487	0.051	0.040	0.538	0.168	0.133
	0.610	0.569	0.403	0.538	0.154	0.041	0.571	0.427	0.259
	0.682	0.789	0.455	0.601	0.481	0.146	0.598	0.507	0.173
	0.765	1.040	0.517	0.654	0.640	0.113	0.667	0.722	0.192
	0.807	1.182	0.457	0.706	0.873	0.144	0.688	0.780	0.064
	0.908	1.476	0.554	0.763	0.962	0.032	0.698	0.813	0.149
	0.932	1.547	0.417	0.830	1.277	0.135	0.715	0.971	0.208
	0.935	1.557	0.214	0.830	1.277	0.090	0.715	0.971	0.414
0.935	1.557	0.000	0.935	1.557	0.000	0.935	1.557	0.000	
Average			0.290			0.072			0.150

b was set to 200, and the unit price c_i is an integer that was randomly assigned a value 1–5, because a

budget that is too large or unit prices that are too small ease the problem by increasing the number

of feasible combinations and thus might be unable to verify the effectiveness of the proposed methods. The experiments were conducted on a PC with a Pentium IV 1.8 GHz CPU, and the average execution time was about 3–5 minutes. For the data of each year, 10 trials with each GA were conducted for each target portfolio. Table 3 shows the average results of each of the 10 trials. The deviation between the obtained and target portfolios are measured in terms of Euclidean distance.

Regardless of the year, the average deviations lead to an obvious and consistent conclusion that Model 5 obtains portfolios closest to the target. Compared with Model 3 and Model 6 obtains portfolios farther from the target than those by Model 3 yet much better than those by Model 2. As stated above, constraint (4) in Model 2 confines the solutions to the area above the horizontal line $r_p = r$ and below the efficient frontier. Compared with other models, the smaller feasible region for Model 2 leads to poorer solutions. Unless the DM insists that the r_p of the obtained portfolio must exceed the desired r , Model 2 does not seem to be a good approach to the portfolio selection problem. Also, Figs. 3–6 show that the GA obtains portfolios very close to the efficient frontiers, indicating that the proposed approach is efficient in finding efficient portfolios under the constraint of minimum transaction lots. Figs. 3 and 4 show that the GAs obtain integer solutions whose (σ, r) are very close to the target values. However, Figs. 5 and 6 reveal that the differences between the integer and target solutions grow large as the fund number increases. Fund number may be the main cause of the solution degeneration because the fund number increases the population and chromosome sizes and thus the solution space such that the GA failed to find efficient integer solutions within 5000 iterations.

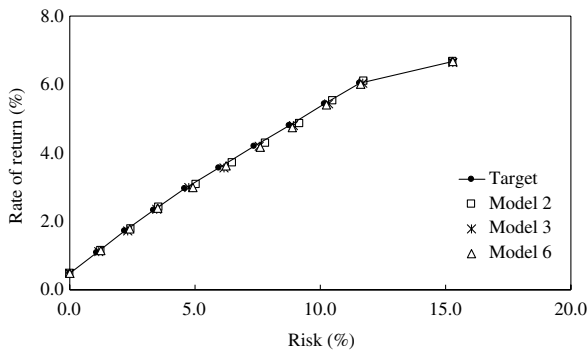


Fig. 3. Model performance with funds in 1997.

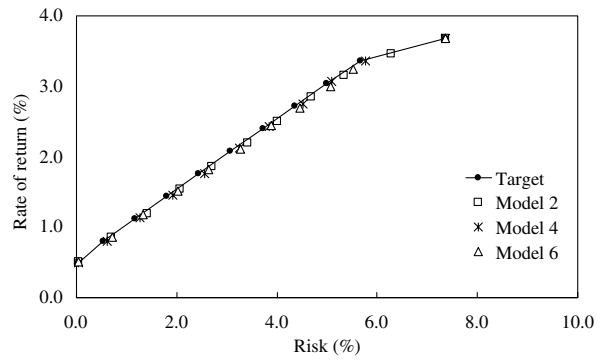


Fig. 4. Model performance with funds in 1998.

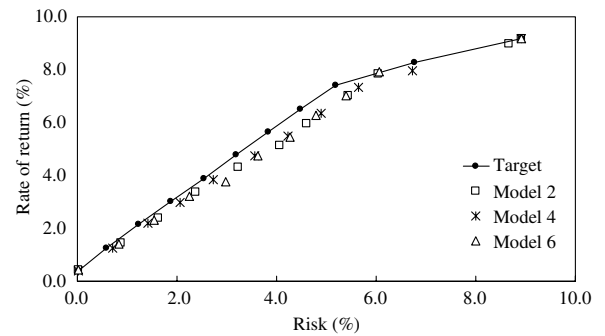


Fig. 5. Model performance with funds in 1999.

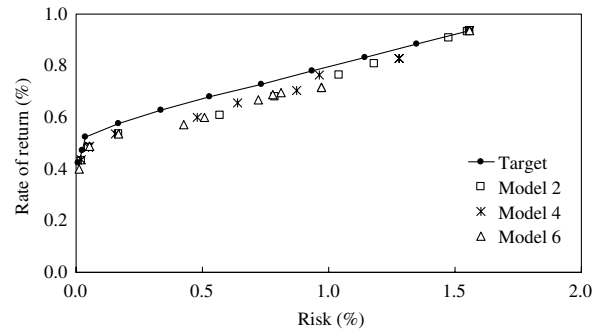


Fig. 6. Model performance with funds in 2000.

Another cause might be that compared with those of 1999 and 2000, more mutual funds of 1997 and 1998 approach their respective efficient frontiers so that more combinations of assets near the efficient frontiers can be found, increasing the possibility of finding better solutions. Consequently, increasing the program running time or reducing the fund number might help when the number of assets is large.

Fig. 7 shows the experimental results obtained by Model 6 for the year 2000 after the fund number is

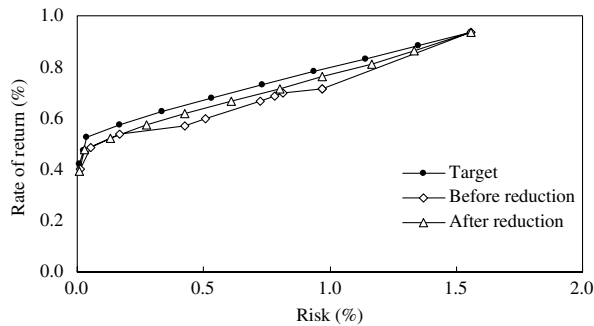


Fig. 7. The effects of reducing asset number.

reduced from 204 to 47 by simply eliminating funds with negative expected rate of return. Though the eliminating method is naïve, the efficient frontier does not seem to change, and the performance was remarkably improved after the asset number was reduced. Nevertheless, the naïve and rough method is not recommended because the efficient frontier might change in other cases. A more sophisticated method remains to be investigated to retain the efficient frontier while reducing the asset number.

According to the experimental results, all the proposed models obtained the solutions within a few minutes, with an average budget utilization rate above 96%. Models 3 and 6 appear more effective than Model 2. However, as Models 2 and 3 depends on the specification of a desired rate of return and additionally requires solving for the minimum risk corresponding to the specified rate of return using Markowitz' model. Besides, the clients need to assure themselves that their desired rates of return are sufficient to reflect their preference for risk. On the other hand, Model 6 (the fuzzy MODM approach) requires only the weights regarding return and risk, which can be applied in various situations and time periods if only the client's preference for risk aversion remains unchanged. Therefore, Model 6 is recommended for practical use. Notably, the number of assets significantly influences the performance of the proposed GAs even if the program execution time is prolonged. An effective way of reducing the asset number without changing the efficient frontier is crucial in finding quality solutions.

6. Conclusion

This study proposes decision models for portfolio selection problems with minimum transaction lots

and uses genetic algorithms to solve the models. The proposed models include one directly modified from the well-known Markowitz' model, one minimizing the distance between the target and obtained portfolios, and one derived from a fuzzy multi-objective decision making approach. Genetic algorithms were devised to solve the portfolio selection problem formulated by the models. The results of empirical studies show that the genetic algorithms for these models can obtain near-optimal within a reasonably short time. The obtained solutions not only are applicable in practice, but also exhibit high mean-variance efficiency. The model incorporates a fuzzy multi-objective decision-making approach is recommended because of its adaptability and simplicity. With this approach, the DM is able to express his/her preferences for the return and risk through assigning the weights of return and risk. Notably, sifting assets ahead can not only save the computation time, but also improve the solution quality. Consequently, how to sift the assets without changing the efficient frontier and the solution quality is not only one of the possible future works, but is critical to solving the portfolio selection problem. Another possible future research direction is how to adjust existing portfolios to look after both transaction costs and the workability and efficiency of portfolios.

Acknowledgements

The authors thank the National Science Council of the Republic of China, Taiwan for financially supporting the research under Contract No. NSC 91-2416-H-168-005.

References

- [1] H. Markowitz, Portfolio selection, *Journal of Finance* 7 (1952) 77–91.
- [2] M.G. Speranza, A heuristic algorithm for a portfolio optimization model applied to the Milan Stock Market, *Computers and Operations Research* 23 (1996) 433–441.
- [3] R. Mansini, M.G. Speranza, On selection a portfolio with fixed costs and minimum transaction lots, Report no. 134, Dip. Metodi Quantitativi, University of Brescia, Italy, 1997.
- [4] R. Mansini, M.G. Speranza, Heuristic algorithms for the portfolio selection problem with minimum transaction lots, *European Journal of Operational Research* 114 (1999) 219–233.
- [5] H. Kellerer, R. Mansini, M.G. Speranza, Selecting portfolios with fixed costs and minimum transaction lots, *Annals of Operations Research* 99 (2000) 287–304.
- [6] H. Konno, H. Yamazaki, Mean-absolute deviation portfolio optimization model and its application to Tokyo Stock Market, *Management Science* 37 (1991) 519–531.

- [7] H. Konno, A. Wijayanayake, Portfolio optimization problem under concave transaction costs and minimal transaction unit constraints, *Mathematical Programming* 89 (2001) 233–250.
- [8] R. Mansini, M.G. Speranza, An exact approach for portfolio selection with transaction costs and rounds, *IIE Transactions* 37 (2005) 919–929.
- [9] J. Shoaf, J.A. Foster, The efficient set GA for stock portfolios, In: *Proceedings of the Decision Science Institute, Orlando, 1996*, pp. 571–573.
- [10] S.M. Lee, A.J. Lerro, Optimizing the portfolio selection for mutual funds, *Journal of Finance* 28 (1973) 1087–2001.
- [11] K. Konno, H. Shirakawa, H. Yamazaki, A mean-absolute deviation-skewness portfolio optimization model, *Annals of Operations Research* 45 (1993) 205–220.
- [12] M. Arenas Parra, A. Bilbao Terol, M.V. Rodríguez Uría, A fuzzy goal programming approach to portfolio selection, *European Journal of Operational Research* 133 (2001) 287–297.
- [13] H.-J. Zimmermann, Fuzzy programming and linear programming with multiple objective functions, *Fuzzy Sets and Systems* 1 (1978) 45–55.
- [14] C.C. Lin, A weighted max–min model for fuzzy goal programming, *Fuzzy Sets and Systems* 142 (2004) 407–420.
- [15] P. Chang, Y. Chen, A fuzzy multicriteria decision making method for technology transfer strategy selection in biotechnology, *Fuzzy Sets and Systems* 63 (1994) 131–139.
- [16] W.F. Abd El-Wahed, M.A. Abo-Sinna, A hybrid fuzzy-goal programming approach to multiple objective decision making problems, *Fuzzy Sets and Systems* 119 (2001) 71–85.
- [17] M.H. Rasmy, S.M. Lee, W.F. Abd El-Wahed, A.M. Ragab, M.M. El-Sherbiny, An expert system for multiobjective decision making application of fuzzy linguistic preferences and goal programming, *Fuzzy Sets and Systems* 127 (2002) 209–220.
- [18] L.A. Zadeh, Fuzzy sets, *Information and Control* 8 (1965) 338–353.
- [19] R. Bellman, L.A. Zadeh, Decision-making in a fuzzy environment, *Management Sciences* 17B (1970) 141–164.
- [20] R. Narasimhan, Goal programming in a fuzzy environment, *Decision Sciences* 11 (1980) 325–338.
- [21] R. Narasimhan, On fuzzy goal programming—some comments, *Decision Sciences* 12 (1981) 532–538.
- [22] E.L. Hannan, Linear programming with multiple fuzzy goals, *Fuzzy Sets and Systems* 6 (1981) 235–248.
- [23] R.N. Tiwari, S. Dharmar, J.R. Rao, Priority structure in fuzzy goal programming, *Fuzzy Sets and Systems* 19 (1986) 251–259.
- [24] D.T. Pham, D. Karaboga, *Intelligent Optimization Techniques: Genetic Algorithms, Tabu Search, Simulated Annealing and Neural Networks*, Springer-Verlag, London, 2000.
- [25] J.H. Holland, *Adaptation in Natural and Artificial Systems*, University of Michigan Press, Ann Arbor, 1975.
- [26] D.E. Goldberg, *Genetic Algorithms in Search, Optimization and Machine Learning*, Addison-Wesley, Reading, MA, 1989.
- [27] K. DeJong, *An Analysis of the Behavior of a Class of Genetic Adaptive Systems*, Ph.D. Thesis, University of Michigan, USA, 1975.
- [28] M. Gen, R. Cheng, *Genetic Algorithms and Engineering Design*, Wiley, New York, 1997.
- [29] H. Levy, Portfolio performance and the investment horizon, *Management Science* 18 (1972) B645–B653.
- [30] S.N. Chen, C.F. Lee, The sampling relationship between Sharpe’s performance measure and its risk proxy: sample size, investment horizon and market conditions, *Management Science* 27 (1981) 607–618.