The Adaptive Genetic Algorithms for Portfolio Selection Problem

Wei-Guo Zhang[1,2] *, *Wei CHEN*[1], *and Ying-Luo Wang*[2]

[1] *School of Business Administration, South China University of Technology, Guangzhou, 510641, P.R. China*
[2] *The School of Management, Xi'an Jiaotong University, Xi'an, Shaan'xi 710049, P. R. China*

Summary
Genetic algorithms (GA) are stochastic search techniques based on the mechanics of natural selection and natural genetics. In this paper, the adaptive genetic algorithms are applied to solve the portfolio selection problem in which there exist both probability constraint on the lowest return rate of portfolio and lower and upper bounds constraints on the investment rates to assets. First, the stochastic model of portfolio selection and it's the reliability decision are presented. Second, the adaptive genetic algorithm to solve the reliability decision is given. Finally, a numerical example of portfolio selection problem is given to illustrate our proposed effective means.
Key words:
Adaptive genetic algorithms, Portfolio selection, stochastic optimization

1. Introduction

John Holland is the founder of the field of genetic algorithms (GAs) ([1]). With the publication of Adaptation in Natural and Artificial Systems in 1975, Holland discussed the ability of simple bit-string representation to encode complicated structures and the power of simple transformations to improve such structures. Genetic algorithms are stochastic adaptive algorithms that the result is a directed random search procedure. The process begins by constructing a random population of possible solutions. This population is used to create a new generation of possible solutions, which is then used to create anther generation of solutions, and so on. The best elements of the current generation are used to create the next generation. It is hoped that the new generation will contain "better" solutions than the previous generation. In this respect, a wide but diverse range of applications ([2], [3]) bear testimony to the fact that genetic algorithms are very useful in solving complicated problems by mimicking some facets of natural evolution. Recently, many applications of genetic algorithm have been applied to the portfolio selection problems ([4], [5]), in which the main research works are related to Markowitz's mean-variance model [7]. Yet, there are some limits and shortcomings in Markowitz's model [7]. In this paper, we propose the stochastic model of portfolio and it's the α -reliability decision, which is different from Markowitz's model and the nation of efficient portfolio. We study the problem of calculating this new model by adaptive genetic algorithm. In addition, a numerical example of a portfolio selection problem is given to illustrate our proposed effective means.

Manuscript revised August 22, 2005.

2. Stochastic Model and Reliability Decision of Portfolio Selection

In Markowitz's mean-variance model, portfolio selection problem with $n(n \geq 2)$ risky assets is usually described as follows. The return rate r_i for asset i is a random variable with expected return $\overline{r_i} = E(r_i), i = 1,2,\Lambda\ ,n.$ Let x_i be the investment rate to asset i. In order to describe conveniently, we introduce the following notations:

$X = (x_1, x_2, \Lambda\ , x_N)^T$, $R = (r_1, r_2, \Lambda\ , r_n)^T$,

$\overline{R} = (\overline{r_1}, \overline{r_2}, \Lambda\ , \overline{r_n})^T$, $F = (1,1,\Lambda\ ,1)^T$, $D = (\sigma_{ij})_{n\times n}$.

$\overline{R}$ and D are the expect return vector and covariance matrix of returns, respectively. Then the return associated with the portfolio $X = (x_1, x_2, \Lambda\ , x_N)^T$ is $r = X^T R$.

The expected return and variance of r are, respectively, given by

$$E(r) = X^T \overline{R}, \quad \sigma^2 = X^T D X .$$

Markowitz's mean-variance model of portfolio selection problem may be described by the following quadratic programming:

$$\begin{cases} \min & \sigma^2 = X^T D X \\ s.t & X^T \overline{R} \geq R_0, \\ & X^T F = 1, \\ & X \geq 0, \end{cases} \tag{1}$$

where R_0 is the lowest expected return of investor.

In practical investment problem, we need to estimate $\overline{R}$ and D. It is well-known fact that the returns of risky assets vary from time to time, and the future states of returns and risks of risky assets cannot be predicted accurately. Moreover, most investor may request the lower and upper bounds constraints on the investment rates to assets. Based on this fact, we propose the stochastic model of portfolio as follows:

$$\begin{cases} \min & \sigma^2 = X^T DX \\ s.t & X^T R \geq R_0, \\ & X^T F = 1, \\ & W \geq X \geq U, \end{cases} \tag{2}$$

where $W = (w_1, \ldots, w_n)^T$ and $U = (u_1, \ldots, u_n)^T$ $(\sum_{i-1}^{n} u_i < 1, 1 \geq w_i \geq u_i \geq 0, \quad i = 1,2,\ldots,n)$ represent the upper and lower bounds constraints on X, respectively.

Due to the constraint $X^T R \geq R_0$ in (2) is a random event. Hence, the feasible solution to (2) may be feasible or not. Both of them exist the degree of the probability, the same to the result of the portfolio decision. Hence, we bring into the measure named reliability and set up a probability-restricted model. The new model can be specified as follows:

$$\begin{cases} \min & \sigma^2 = X^T DX \\ s.t & P(X^T R \geq R_0) \geq \alpha, \\ & X^T F = 1, \\ & W \geq X \geq U, \end{cases} \tag{3}$$

where $W = (w_1, \ldots, w_n)^T$ and $U = (u_1, \ldots, u_n)^T$ $(\sum_{i-1}^{n} u_i < 1, 1 \geq w_i \geq u_i \geq 0, \quad i = 1,2,\ldots,n)$ represent the upper and lower bounds constraints on X, respectively.

(3) is called the α-reliability decision model of (2), and the feasible solution to (3) is called the α-reliability feasible solution of (2). So the optimal solution to (3) is a α-reliability decision for the portfolio. The α-reliability decision has more significance because it shows that the portfolio selection is stochastic decision, its' result has two sides: reliability and unreliability. Because of avoiding regarding it as a determinate decision it is better to reflect the uncertain economic environment. Hence, the α-reliability decision is very practical and effective in real investment management.

The model (3) can be transformed into the determinate decision model. If constant M is given by the following formula:

$$P\left(\frac{X^T R - X^T \overline{R}}{\sqrt{X^T DX}} \geq M\right) = \alpha,$$

then the probability constraint condition $P(X^T R \geq R_0) \geq \alpha$ is equivalent to the determinate constraint condition $X^T \overline{R} + M\sqrt{X^T DX} \geq R_0$.

The proof of this conclusion is following:

If $P(X^T R \geq R_0) \geq \alpha$,

then $P(X^T R \leq R_0) \leq 1 - \alpha$.

Since $P(X^T R \geq X^T \overline{R} + M\sqrt{X^T DX}) = \alpha$,

$P(X^T R \leq X^T \overline{R} + M\sqrt{X^T DX}) = 1 - \alpha$.

According to the monotonic increasing of the distribution function, we obtain

$$X^T \overline{R} + M\sqrt{X^T DX} \geq R_0.$$

Contrarily, if $X^T \overline{R} + M\sqrt{X^T DX} \geq R_0$, then

$$P(X^T R \geq R_0) \geq P(X^T R \geq X^T \overline{R} + M\sqrt{X^T DX}) = \alpha$$

So, (3) is equivalent to the following determinate constraint model:

$$\begin{cases} \min & \sigma^2 = X^T DX \\ s.t. & X^T \overline{R} + M\sqrt{X^T DX} \geq R_0, \\ & X^T F = 1, \\ & W \geq X \geq U, \end{cases} \tag{4}$$

where $W = (w_1, \ldots, w_n)^T$ and $U = (u_1, \ldots, u_n)^T$ $(\sum_{i-1}^{n} u_i < 1, 1 \geq w_i \geq u_i \geq 0, \quad i = 1,2,\ldots,n)$ represent the upper and lower bounds constraints on X, respectively.

In (4), M depends on the probability distribution of $X^T R$. If the return rate of every risky asset i follows normal distribution $N(\overline{r_i}, \sigma_i^2), (i = 1,2,\ldots,n)$, then $X^T R$ follows normal distribution $N(X^T \overline{R}, X^T DX)$. Thus, M is determined the following formula:

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{M} e^{-\frac{x^2}{2}} dx = 1 - \alpha$$

From the normal distribution chart $\varnothing(M) = 1 - \alpha$, we can obtain M.

If $M = 0$, $W = (1, K, 1)^T$ and $U = (0, K, 0)^T$, then (4) is equivalent to Markowitz's mean-variance model (1). It is obvious that the α -reliability decision model (3), i.e., (4) is extensions of previous models for portfolio selection problem, such as Markowitz's mean-variance model. Furthermore, previous researches have shown that finding the analytic form solution to (4) is very difficult. In next section we will give the adaptive genetic algorithm for finding the α -reliability decision.

3. The Adaptive Genetic Algorithm

If the multimode function wants to keep the global search ability it must have balanced search ability. Crossover probability p_c and mutation probability p_m are the main factors in affecting balanced search ability (global search ability and local search ability). While we strengthen one ability by increasing or decreasing p_c, p_m, we may weaken other abilities. Both p_c and p_m in the simple genetic algorithm (SGA) are invariant, so for the complex optimal problem the GA's efficiency is not high. In addition, immature convergence maybe caused. Therefore, the goals with adaptive probabilities of crossover and mutation are to maintain the genetic diversity in the population and prevent the genetic algorithms to converge prematurely to local minima. Strinvivas([6]) put forward the adaptive genetic algorithm, and its basic idea is to adjust p_c and p_m according to the individual fitness. This algorithm can better solve the problem of adjusting p_c and p_m dynamically and also fits to all kinds of optimal problem. Based on these facts, we adopt the adaptive genetic algorithm to obtain the optimal solution to (4) as follows:

3.1. Initialization

Define integer H is the initial random population of chromosomes, and we adopt real code. Every chromosome include n gene bit (represent n assets), the gene value is the proportion to this asset in the portfolio. Obviously, the feasible set of (4) includes the following set:

$$\Omega = \{(x_1, \Lambda, x_n) | u_1 \le x_1 \le w_1, \Lambda, u_n \le x_n \le w_n\}$$

It produces random number from Ω and tests its feasibility. If it is feasible then it is a member of the initial population, otherwise, we go on produce random number from Ω until obtain feasible solution. After finite sample there are H initial feasible chromosomes V_1, Λ, V_H.

3.2. Evaluation function

Through evaluation function $eval(V)$ we set probability for every chromosome V, so the selection probability is proportion to their fitness. That is, by roulette wheel selection, chromosomes with a high fitness value have a great chance of being selected to generate children for the next generation., then the chromosome is reallocated by the sequence number instead of target value, and chromosome is arrayed from good to bad. That is say, a chromosome is better sequence number is lower. So we set $\alpha \in (0,1)$ and define evaluating function based on the sequence number

$$eval(V_j) = \alpha(1-\alpha)^{j-1}, \quad j = 1, 2, \Lambda, H.$$

3.3. Selection operator

3.3.1. According to the rule that the selection operator chooses individuals with a probability that corresponds to the relative fitness. Chromosomes with a high fitness value have a great chance of being selected to generate children for the next generation., two chosen individuals, called the parents. We define reproduction probability for $v_j(k)$:

$$p_j(k) = \frac{F(v_j(k))}{F(v_1(k)) + F(v_2(k)) + \Lambda + F(v_H(k))}.$$

$j = 1$ means chromosome is the best one, $j = H$ means chromosome is the worst one.

3.3.2. For $v_j(k)$, $j = 1, 2, \Lambda, H$ calculate cumulative probability q_j:

$$\begin{cases} q_0 = 0, \\ q_j = \sum_{i=1}^{j} p_i(k), \qquad j = 1, 2, \Lambda, H. \end{cases}$$

3.3.3. Produce random number r in $(0, q_H)$. If $q_{j-1} < r < q_j$, then we select the chromosome $v_j(k)$, $j = 1, 2, \Lambda, H$.

3.3.4. Repeating 3.3.2 and 3.3.3 H times, we can obtain H copied chromosomes, defined by $v'(k) = (v_1'(k), v_2'(k), \Lambda, v_H'(k))$.

3.4. Crossover operator

3.4.1. Instead of using fixed p_c, we adjust it adaptively based on the following formula:

$$p_c = \begin{cases} p_{c1} - \dfrac{(p_{c1} - p_{c2})(f' - f_{avg})}{f_{\max} - f_{avg}} & f' \ge f_{avg}, \\ p_{c1} & f' < f_{avg}, \end{cases}.$$

where $f_{\max}$ is the highest fitness value in the population; f_{avg} is the average fitness value in every population; f' is higher fitness value between two individuals; in addition;

We set $p_{c1} = 0.9, p_{c2} = 0.6$.

3.4.2. $v_j'(k)$ is the parent of assured crossover operator, repeat the following step from $j=1$ to $j=k$: produce random number r in $[0,1]$. If $r < p_c$, then we select $v_j'(k)$ as parent.

3.4.3. $v_1''(k), v_2''(k), \Lambda, v_L''(k)$ are selected parent and they are randomly assigned into groups such as $(v_1''(k), v_2''(k)), (v_3''(k), v_4''(k)), \Lambda$, then conduct crossover operator.

3.4.4. Conduct crossover operator for $(v_1''(k), v_2''(k))$ and then do as following:

Produce a random number c in $(0,1)$. Let

$$X_1(k) = cv_1''(k) + (1-c)v_2''(k)$$

$$X_2(k) = (1-c)v_1''(k) + cv_2''(k)$$

If the new chromosomes don't satisfy restraint conditions we refuse it as offspring, then repeat 3.4.4 until we obtain the new chromosomes that are feasible.

3.4.5. Using the same way as above, we conduct crossover operator for other groups.

3.5. Mutation operator

3.5.1. Instead of using fixed p_m, we adjust it adaptively based on the following formula:

$$p_m = \begin{cases} p_{m1} - \dfrac{(p_{m1} - p_{m2})(f - f_{avg})}{f_{\max} - f_{avg}} & f \ge f_{avg} \\ p_{m1} & f < f_{avg} \end{cases}$$

where $f_{\max}$ is the highest fitness value in the population; f_{avg} is the average fitness value in every population; f is the mutation individual fitness; and we set $p_{m1} = 0.1, p_{m2} = 0.001$

3.5.2. Crossover takes two selected parents (chromosomes), conduct the following mutation operator repeatedly from $i=1$ to L: produce random number r in $[0,1]$. If $r < p_m$, then we select $X_i(k)$ as parent of the mutation.

3.5.3. Produce random integer i in $[0,1]$, j in $[0,1]$, and random numbers r_1 and r_2 in $[0,1]$ and then creates offspring by the following method: for the chromosome a_i its i gene bit a_{ii} is replaced by r_1, at the same time, for the chromosome a_i its j gene bit a_{ij} is replaced by r_2. If the new chromosomes don't satisfy restraint conditions we refuse it as offspring, then repeat 3.5.3 until we obtain the new chromosome is feasible. After finite sample we can produce number new mutation individual.

3.5.4. From mutation operator 3.5.3 calculate s new individuals' fitness values; calculate $L-s$ new individuals' fitness values, that new individuals can conduct crossover operator in 3.4 and cannot conduct mutation operator in 3.5; At last, from these new individuals together with $H-L$ unselected individuals in 3.3 we generate new population:

$$v(k+1) = \{v_1(k+1), v_2(k+1), \Lambda, v_H(k+1)\}$$

3.6. Convergence conditions

Stopping test is judged by calculating $|F_i - F_{i+1}| < \varepsilon$ where F_i and F_{i+1} are continuous generation's fitness values, ε is fixed arbitrary decimal fraction. If the results satisfy $|F_i - F_{i+1}| < \varepsilon$ the genetic operator will be stopped.

4. Experimental Result

In this section we provide experimental results to demonstrate the effectiveness of the adaptive genetic algorithm to the stochastic portfolio model. Considering portfolio selection with 9 risky assets, the return rate and covariance chart of returns are shown in Table 1. We make the following assumption in this experiment:

The expected portfolio return rate $R_0 = 10\%$;

$U = (0.05, 0.05, 0.10, 0.13, 0.18, 0.05, 0.10, 0.13, 0.05)^T$,

$W = (0.15, 0.15, 0.20, 0.25, 0.30, 0.15, 0.25, 0.25, 0.15)^T$

$\alpha = 0.70$, $M = -0.53$, population size $H = 30$, $\varepsilon = 0.001$.

At last, we can conclude the investment vector of the optimal portfolio with 70%-reliability:

$$X^* = (0.057, 0.059, 0.112, 0.156, 0.235, 0.052, 0.111, 0.159, 0.059)^T.$$

It holds that $X^{*T}\overline{R} + M\sqrt{X^{*T}DX^*} \ge 0.10$, that is,

$$P(X^{*T}R \ge 0.1) \ge P(X^{*T}R \ge X^{*T}\overline{R} + M\sqrt{X^{*T}DX^*}) = 0.70$$

Table 1

$\overline{r_i}$ %	σ_{ij} %								
7.1	2.0	1.6	1.9	3.3	1.0	2.7	1.8	2.8	2.1
10.1	1.6	3.1	1.7	1.9	0.8	0.7	0.9	2.1	1.5
19.6	1.9	1.7	6.9	6.3	4.5	0.7	0.4	7.8	3.1
20.6	3.3	1.9	6.3	8.0	4.5	2.1	0.9	7.9	2.0
23.4	1.0	0.8	4.5	4.5	8.7	0.6	1.1	8.6	2.9
15.3	2.7	0.7	0.7	2.1	0.6	4.2	1.0	2.0	1.1
13.6	1.8	0.9	0.4	0.9	1.1	1.0	3.6	1.8	0.8
21.6	2.8	2.1	7.8	7.9	8.6	2.0	1.8	8.3	3.5
14.4	2.1	1.5	3.1	2.0	2.9	1.1	0.8	3.5	4.0

5. Conclusion

In this paper, we have discussed the portfolio selection in which there exist both probability constraint on the lowest return rate of portfolio and lower and upper bounds constraints on the investment rates to assets. We have proposed the stochastic portfolio model and it's the reliability decision of portfolio selection, them are extensions of Markowitz's mean-variance model and the efficient portfolio. Particularly, the adaptive genetic algorithm has been applied to obtain the reliability decision of portfolio selection. The numerical result has showed that its application in portfolio selection is reliable and useful.

Acknowledgment

This research is supported by the National Natural Science Foundation of China (No.70571024) and China Postdoctoral Science Foundation (No.2005037241).

References

[1] J. H. Holland, Adaptation in Natural and Artificial Systems. University of Michigan Press. Ann Arbor, 1975.
[2] W. E. Hart, The role of development in genetic algorithms. In D. Whitley and M. Vose (Eds.), Foundations of Genetic Algorithms 3. Morgan Kaufmann. 1994.
[3] David B.Fogel. "An Introduction to Simulated Evolutionary Optimization". IEEE Trans Neural Networks, 1994, Jan.5(1).
[4] Yusen Xia, Baoding Liu, Shouyang Wang, K.K Lai.A new model for portfolio selection with order of expected returns. Computers and Operations Research 2000, 27:409-22.
[5] Yusen Xia, Shouyang Wang, Xiaotie Deng. A compromise solution to mutual funds portfolio selection with transaction costs. European Journal of Operations Research, 2001,134:564-581.
[6] Strinivas M , Patnaik L M . Adaptive Probabilities of Crossover and Mutation In Genetic Algorithms. IEEE Trans.Syst. Man and Cybernetics , 1994,24(4):656-667.
[7] H. Markowitz, Analysis in portfolio choice and capital markets, Oxford, Basil Blackwell, 1987.

Wei-Guo Zhang received the B.E. degree from Ningxia Univ. in 1983, M. E. degree from Sichuan Univ. in 1995 and Dr. degree from Xi'an Jiaotong Univ. in 2003, respectively. After working as a research assistant (from 1985), an assistant professor (from 1990) and an associate professor (from 1995) in the Dept. of Math. and Electrical Engineering, Ningxia Univ., he had been a professor at Ningxia Univ., from 1998 to 2004. Now, he has been a professor at South China University of Technology, since 2004. His research interest includes information processing, artifial intelligences, design and analysis of algorithms and their application to portfolio managements.