

Research on the Optimal Portfolio Based on Genetic Algorithms

Jun Han

Glorious Sun School of Business and Management
Donghua University
Shanghai, China, 200051
e-mail: hj_87421@sina.com

Abstract—Investment in securities is in an uncertain environment, any gains obtained are accompanied by certain risks. The essence of portfolio optimization is the optimal allocation of the limited assets in securities with different risk and return characteristics. In this paper, the portfolio decision-making utility function is established based on E-SV risk measure, through the analysis of Markowitz portfolio model, an improved portfolio selection criterion is obtained. Because what we solve is a more complex fractional programming portfolio selection model, the traditional algorithm can not guarantee to get global optimum. In view of this, the genetic algorithms of random simulation are introduced to conduct an in-depth study. Empirical analysis shows that the portfolio model and algorithm proposed in this paper is scientific and reasonable, which may provide investors with an effective theoretical guidance and basis for decision making.

Keywords-Genetic Algorithms; Portfolio; Utility Function

I. INTRODUCTION

Stock market is a high risk market, which is subject to national economic situation, policy changes and the laws of the market itself and other factors, and it is filled with many unpredictable factors. How to avoid the risk of stock market investment is becoming an eternal topic for these investment theorists and practitioners. Therefore, investors constantly seek some strategies which can make investments in low-risk and high return. The existence of the portfolio makes each invested securities properly decentralized; through the offset of a variety of security risks, making per unit revenue risk the smallest or per unit yield the largest.

Many experts and scholars carried out extensive research on this question. In 1952, the selection concept of the portfolio put forward by the famous American economist Markowitz marks the beginning of the modern theory of the securities. The paper believed that a rational investor not only hope "high rate of return," but also hope "to determine possible gains." This means that investors seek "to maximize the expected return," at the same time, they pursue the "minimal uncertainty." In the initial stage of investment decision-making, it is needed to strive to make the goal of these two mutual restraints for a balance, this paper proposes a classic mean - variance model, which respectively uses the

expected return and variance of returns to measure the level of the expected return on investment and uncertainty (risk). This theory makes the risk to be quantified for the first time, and its mystery gradually fade, Finance began to get rid of purely descriptive studies and the state of thumb operation, quantitative methods come into the field of financial research. Since then, the economists have been using the quantitative methods to continuously improve portfolio management theory and the effective management approach of actual investment, making the portfolio theory have a significant progress in the innovation of basic concepts, improvement of the theoretical system, the demonstration of the important theory and expansion of the theoretical application.

In recent decades, with the cross-cutting penetration and promotion of life science and engineering science, forming a kind of random search algorithm in which machines simulate natural processes to solve complex problems - evolutionary algorithm. As one of the evolutionary algorithms, genetic algorithms are based on Darwin's theory of biological evolution to simulate of natural biological evolution process, and get the best results through survival of the fittest. Compared with the traditional optimization method, genetic algorithms have strong robustness and high efficiency, and can get the global optimal solution with the greater the probability. Currently, genetic algorithms have been applied to many areas of scientific research and engineering and have made great success in solving various combinatorial optimization searching and optimization calculation problems in scientific research and engineering technology. Therefore, genetic algorithms are applied to solve the portfolio problem, which has important academic and practical significance in economics, and at the same time, it also has important practical significance for China's economic development.

II. G A (GENETIC ALGORITHM)

Genetic Algorithms (Genetic Algorithms - GAs) are a random search optimization algorithm formally proposed by Holland, an American professor of the University of Michigan in 1975. It is a genetic selection and natural selection process mainly through the simulation of biological evolution, and gets the best one or more offspring. Genetic algorithms were applied in the field of financial investment beginning in the 90 s, compared with the traditional optimization methods, genetic algorithms adopted many unique methods and technology, making the genetic

algorithms have the following advantages in computerized on the analysis of stock investment technology:

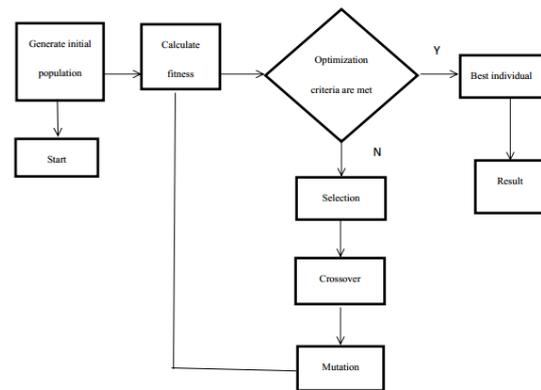
A. In technical analysis, especially for the subjective technical analysis, type, the graph is one of the most basic tools and handling objects. It is required that the technical analysis computerized technology should have a strong graphics processing capability. Many optimization methods just do not have the capability, or they lack the capacity. The processing target of genetic algorithms is not the parameters to be optimized themselves, but the individual to code the parameter sets. Through directly processing the graphics, relationships, etc which are not easy to be handled by the computer into numerical coding, genetic algorithms can easily operate and optimize them.

B. For rational investors, the stock investment decisions are the results of weighing the risks and benefits according to their own investment philosophy on a number of options. There is never the optimal solution in the stock investment decisions we all agree. Genetic algorithms simultaneously process and optimize multiple individual in the groups and get several more desirable solutions. Therefore, compared with the optimization method to obtain a single optimal solution, genetic algorithms which can find several excellent alternatives are undoubtedly better meet the actual needs of investors.

C. Stock investment environment is variable with many influencing factors. No absolute things are in stock market, the occurrence of any event is possible, just the probability is different. And traditional optimization methods use deterministic rules, their optimizing process and the results are often determined. Genetic algorithms use probability transition rules to guide its search direction. This not only expands the range of possible solutions, but also makes the solution more representative and better meet the requirements of investors.

D. Although many factors affect the stock market, but the technical analysis focuses only on the stock price. The information requirements of genetic algorithms for background and expertise to be solved are rarely, it only evaluate the merit of the solution through the fitness, to guide the optimization process. Thus it is suitable for technical analysis.

The most unique feature of genetic algorithms is three kinds of genetic operators: selection, crossover and mutation. Through the close cooperation and joint action of these three kinds of genetic algorithms operation, genetic algorithms guide the population continually to evolve in the right direction, and get a satisfactory solution. Its basic flow is as follows:



Genetic operation is conducted for numerous individuals at the same time. The numerous individuals form a group, the first generation of group is known as the initial population, individual in the initial population is produced by random method. In the process of genetic algorithms, as a starting point, initial group evolves until the evolution process stops evolution according to some kind of stopping criteria, resulting in the last generation. The number of individuals in the group is known as the population size, in practice, it is needed to keep moderate population scale. Genetic algorithms basically don't use the external information in the evolutionary search, only using evaluation function, namely fitness function to evaluate the merits of the individual or solution, and is used as the basis of genetic operation later. In specific issues, the design of fitness function is set according to the requirements of solving problem itself; usually the target function is used as fitness function.

Genetic operation is to simulate the operation of bio-genetic, its task is to evaluate the individual's fitness of group and exert a certain operation according to the fitness, realizing the evolutionary process of survival of the fittest. From the perspective of optimization search, genetic operation can optimize solution of the problem from generation to generation, and approximate the optimal solution. The operation of selecting superior individual and eliminating inferior individual from the population is called selection. Selection operation is based on individual fitness evaluation; the selection purpose is to directly inherit the optimized individual to the next generation or produce new individual through matching cross to the next generation. The commonly used selection methods have fitness ratio method, the best individual preservation method. Crossover plays a central role in genetic operation; it refers to the operation of generating new individuals from part of the structure of two parent individuals to be replaced and restructured, through the cross, the search capability of genetic algorithm is improved greatly. As far as binary encoding is concerned, the crossover operation sets intersection for the randomly paired individuals according to preset crossover probability P_c , and then exchange parts of the structure of the of individual paired before and after these points and generate two new individuals. Mutation refers to the pre-set mutation probability P_m to change the genetic

value of certain gene locus of the individual string. As far as binary encoding is concerned, mutation operation is to take the reverse genetic value, namely, 0→1 or 1→0. The main purpose of GA introducing mutation is to enhance the local search ability and to maintain the diversity of population.

In general, the termination condition of the genetic operation is as follows: If the individual evolution is found to be in a stabilized state, namely the individual accounting for a certain proportion in a certain generation is found to be entirely the same individual, which can be considered to have found the optimal solution.

III. GENETIC ALGORITHMS ARE USED TO SOLVE THE PORTFOLIO.

A portfolio P is composed of six risk securities, the covariance between the returns and returns is shown in the following table:

	Return	Covariance					
		(%)	1	2	3	4	5
1	18.5	0.210	0.210	0.221	-0.216	0.162	-0.215
2	20.3	0.210	0.225	0.239	-0.216	0.168	-0.219
3	22.9	0.210	0.239	0.275	-0.246	0.189	-0.247
4	21.8	-0.216	-0.216	-0.246	0.256	-0.185	0.254
5	16.7	0.162	0.618	0.256	-0.185	0.142	-0.188
6	23.1	-0.215	-0.219	-0.185	0.254	-0.188	0.266

The issue needed to be considered is under the condition of a given portfolio expected return, how to use genetic algorithms to solve combination P, making the risk minimum.

A. Initialization of the population

Standard genetic algorithms are based on a binary string as a coding way of chromosome, but as far as the large-scale multi-variable optimization is concerned, the decimal coding has the characteristics of faster computing, more accurate and easier to design joined problem domain knowledge of the genetic operator compared with the binary coding, which is very important for processing constrained optimization problem such as securities portfolio models by using genetic algorithms. Randomly generate m initial population with a uniform distribution, each individual

$x' = (x'_1, x'_2, \dots, x'_n)^T$, $x'_i \in (0,1)$ is expressed as a chromosomal gene encoding, namely the investment ratio of various securities, and it is processed by the normalization process, that is, $x_i = x'_i / (\sum_{i=1}^n x'_i)$,

s.t. $\sum_{i=1}^n x_i = 1$, the initial solution is obtained.

B. Calculation of fitness value

In 1952, Markowitz proposed portfolio investment mean - variance model, but Markowitz model has also deficiencies

in practical applications. In view of the deficiencies in Markowitz model, semi-variance of securities return rate is used as a measure of securities investment risk (the $E - SV$ risk measure), for a portfolio, its risk measure is as follows:

$$f^-(x) = \sum_{i=1}^n \sum_{j=1}^n \sigma_{ij}^- x_i x_j$$

In which, $\sigma_{ij}^- = E(R_i - E(R_i))^- (R_j - E(R_j))^-$, ($i, j = 1, \dots, n$), R_i is the return of security i , in the $E - SV$ risk measure, only when the actual rate of return is lower than expected return, do securities generate investment risk, and when the actual rate of return is higher than the expected rate of return, there is no risk. So the E-SV risk measure can better reflect the risk of securities investments.

Later, some scholar constructed the optimal risk objective function,

$$f(x) = \sum_{i=1}^n \sum_{j=1}^n \sigma_{ij}^- x_i x_j - \sum_{i=1}^n \sum_{j=1}^n \sigma_{ij}^+ x_i x_j,$$

$$\sigma_{ij}^- = E(R_i - E(R_i))^- (R_j - E(R_j))^- , (i, j = 1, \dots, n),$$

$$\sigma_{ij}^+ = E(R_i - E(R_i))^+ (R_j - E(R_j))^+ , (i, j = 1, \dots, n),$$

based on $f^-(x)$,

In the E-SV risk measure, the optimal risk objective function is

$$f(x) = \sum_{i=1}^n \sum_{j=1}^n \sigma_{ij}^- x_i x_j - \sum_{i=1}^n \sum_{j=1}^n \sigma_{ij}^+ x_i x_j, \text{ when it is}$$

minimized, on one hand, it minimizes $\sum_{i=1}^n \sum_{j=1}^n \sigma_{ij}^- x_i x_j$,

on the other hand, it maximize $\sum_{i=1}^n \sum_{j=1}^n \sigma_{ij}^+ x_i x_j$. Because

$\sum_{i=1}^n \sum_{j=1}^n \sigma_{ij}^+ x_i x_j$ reflects the discrete degree of the expected return higher than the actual return. The larger $\sum_{i=1}^n \sum_{j=1}^n \sigma_{ij}^+ x_i x_j$ is, the greater the portfolio return is. Then risk utility function is defined as follows:

$$g(x) = \frac{\sum_{i=1}^n \sum_{j=1}^n \sigma_{ij}^- x_i x_j}{\sum_{i=1}^n \sum_{j=1}^n \sigma_{ij}^+ x_i x_j}$$

The utility function $g(x)$, is used to establish the following portfolio investment decision model.

$$\min g(x) = \frac{\sum_{i=1}^n \sum_{j=1}^n \sigma_{ij}^- X_i X_j}{\sum_{i=1}^n \sum_{j=1}^n \sigma_{ij}^+ X_i X_j}$$

$$s.t. \quad \sum_{j=1}^n r_j X_j \geq r_0$$

$$0 \leq x_j \leq \delta_j, \quad j = 1, \dots, n$$

In which, $r_i (i=1, \dots, n)$ refers to the investment return of Risk i , $x_i (i=1, \dots, n)$ refers to the investment ratio of Risk bond i .

The objective function of the problem to solve is regarded as the evaluation criteria of genetic operation, namely the fitness function $\min g(x)$.

C. Selection

Roulette Wheel Selection is used, this selection strategy is commonly used in the genetic algorithms, which firstly calculates the relative fitness of the individual $\frac{f_i}{\sum_{i=1}^N f_i}$,

denoted as p_i , then divide a disk into n parts according to the probability of selection $\{p_i, i=1, 2, \dots, n\}$, wherein the central angle of Sector i is $2\pi p_i$. In selecting, suppose you can rotate the disc, if a reference point falls into Sector i , then select Individual i , this selection strategy can be implemented as follows: firstly, generate a random number r in the interval of $[0,1]$, if $P_1 + P_2 + \dots + P_{i-1} < r \leq P_1 + P_2 + \dots + P_i$, then Individual i is selected, obviously, this selection method is quite similar to the wheel of the roulette wheel selection, the larger the area of the sector is, the greater the probability of the sieve falling into the sector is, that is to say, the greater the fitness of the individual, the greater the chance to be selected, thus the greater the possibility of its genetic structure inherited to the next generation.

The Roulette Wheel Selection method is adopted here; its selection process is described as below:

Step 1: Compute the cumulative probability a_p for each chromosome $V_p (p=1, 2, \dots, m)$;

Step 2: Generate the random number r in the interval of $[0,1]$;

Step 3: if $r \leq a_1$, select the first chromosome V_1 ; otherwise select the P th chromosome V_p which meets the requirement $a_{p-1} < r < a_p$;

Step 4: Repeat steps 2 and 3 until the chromosome $m-1$ is obtained.

D. Crossover

Generate the random number r in the interval of $[0,1]$; if $r < P_c$, in which, P_c is crossover probability, then select the current chromosome to cross, repeat the process m times, each chromosome (V_1, V_2) is crossed to produce the next two generation (V'_1, V'_2) ; $V'_1 = c_1 V_1 + c_2 V_2$; $V'_2 = c_1 V_2 + c_2 V_1$; wherein, $c_1 + c_2 = 1$, in which, c_1 is a random number in the interval of $[0,1]$.

E. Mutation

Uniform Mutation is adopted here, mutation probability P_m defines the number of the expected chromosome of mutation operation, generate a random number r in the interval of $[0,1]$ of each chromosome, if $r < P_m$, then the selected chromosome is mutated, for the selected offspring, select its elements to mutate, the generated offspring is $V' = [x_1, \dots, x'_i, \dots, x_n]$, in which, x'_i is in the interval of $[x_i^L, x_i^U]$.

F. Result

In the above genetic algorithms operation, suppose $P_c = 0.6, P_m = 0.005$, Matlab is adopted to program and calculate, the optimal solution $x^* = (0.1688 \ 0.2311 \ 0.1839 \ 0.1219 \ 0.0883 \ 0.206)$ is obtained. The six kinds of securities portfolio obtained achieve the expected utility of the investment risk $g(x) = 0.0220$ in the non-negative investment ratio constraint conditions.

IV. CONCLUSION

It can be seen through the examples:

Genetic algorithms can obtain a very good effect for solving the portfolio problems containing integer programming, can accurately give the proportion of securities in the portfolio, and get the returns and risks corresponding to the portfolio. But in using genetic algorithms to solve the specific optimization problems, it is needed to be combined with the knowledge related to the problem solving field to redesign the algorithms so as to achieve better effect.

REFERENCES

- [1] Markowitz H M. Portfolio Selection[J]. Journal of Finance, 1952, 7: 77-91
- [2] Tanaka H, Peijun G, Trksen I B. Portfolio selection based on fuzzy probabilities and possibility distributions. Fuzzy Sets and Systems. 2000, 111: 387-397
- [3] Patel N R, Subrahmanyam Marti G. A simple algorithm for optimal portfolio selection with fixed transaction costs. Management Science. 1982, 28: 303-314.
- [4] Chang T J, Meade N, Beasley J E, et al. Heuristics for cardinality constrained portfolio optimization. Computers and Operations Research, 2000, 27(13): 1271-1302

- [5] Kellerer H, Mansini R, Speranza M G. On selecting a portfolio with fixed costs and minimum lots. *Annals of Operations Research*. 2000,99(3):287-304
- [6] Chen Wei. Research on Portfolio Selection Model and Heuristic Algorithm [Doctoral Dissertation]. Beijing: Management Science and Engineering of Beijing Jiaotong University, 2007-03
- [7] Wang Yanqing. Research on Portfolio in the Probability Criterion and Fuzzy Criteria. [Doctoral Dissertation]. Tianjin: Management Science and Engineering of Tianjin University. 2006 — 05
- [8] Ma Yongkai. Research on Portfolio Investment Decision-Making Method Based on Factor Model. [Doctoral Dissertation]. Xi'an: University of Electronic Science and Technology, Management. 2005-04
- [9] Guo Zhigang. Securities Investment Analysis Based on Neural Networks and Genetic Algorithms. [Doctoral Dissertation]. Sichuan: Southwest University of Finance and Economics Finance. 2006-04
- [10] Wu Guoyun. Optimization Model and Algorithm of Portfolio. [Master Dissertation]. Beijing: Beijing University of Chemical Technology, Applied Mathematics. 2007—05
- [11] Qin Changcheng. Optimal Solution Algorithm and Parameter Analysis of An Portfolio Model. [Master Dissertation] Henan, Henan University, Applied Mathematics. 2006-05