Genetic Algorithm for Designing an Index Fund

02991784  大阪大学経済学研究科  *韓 尚憲  HAN Sang-Heon
01300234  大阪大学経済学研究科  田畑吉雄  TABATA Yoshio

1  INTRODUCTION

It is well known that a stock index of a market portfolio, such as TOPIX, S & P 500 etc., is a measure of the performance of a hypothetical basket of stocks. In recent, it has become very popular in the investment industry to design portfolios or an index fund with a securities that closely track visible indices.

In this study, we consider the traditional asset allocation of Markowitz type and develop an efficient method based on GA (Genetic Algorithm) for designing an index fund. The GA is known one of the most powerful tools for solving hard constrained combinatorial problems such as Knapsack Problem, Travelling Salesman Problem, Portfolio Selection Problem, etc. The problem of minimizing the tracking error under the given number of securities in the index fund is a combinatorial problem, which is formulated as quadratic programming problem with 0-1 variables[1]. Since it is very difficult to solve in practice, we propose an efficient algorithm. Next, we discuss the efficiency of the algorithm by utilizing the Japanese index Nikkei 225 and demonstrate several modifications of the basic genetic procedures including a new fitness and mutation rate.

2  MODEL

We are particulary interested in forming an index-like m security portfolio $R_m$ of the n original securities (n > m) in the mean-variance framework of Markowitz, satisfying the traditonal criterion that the tracking error $E(R_n - R_m)^2$ is minimized. It is varified that tracking error between portfolios $R_n$ and $R_m$ is

$$E(R_n - R_m)^2 = Var(R_n - R_m) + (E(R_n) - E(R_m))^2.$$ 

Let $y_i (i = 1, 2, \ldots, n)$ be a 0-1-variable defined as

$$y_i = \begin{cases} 
0, & \text{if the } i\text{th security is not included in } R_m \\
1, & \text{if the } i\text{th security is included in } R_m.
\end{cases}$$

Then the problem to minimize tracking error can be written as

$$\text{Minimize } E(R_n - R_m)^2 = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij}(w_i - x_iy_i)(w_j - x_jy_j)$$

subject to

$$x_1y_1 + x_2y_2 + \cdots + x_ny_n = 1$$

$$y_1 + y_2 + \cdots + y_n = m$$

$$y_i = 0 \text{ or } 1.$$ 

$$R_n = \sum_{i=1}^{n} w_iP_i \quad (R_m = \sum_{i=1}^{n} x_iy_iP_i):$$

random return of n(m) portfolios

$a_{ij}$: variance-covariance matrix of security $i$ and $j$.

$w_i$: fraction of investor's wealth invested in security $i$.

$x_i$: weight of security $i$ ($i = 1, 2, \ldots, n$)

This problem is a quadratic programming problem with 0-1 variables. However this optimization problem is very difficult to solve, since it belongs to the class of NP (non-polynomial) complete problems, even if the values of $x_i$ are given. The GA approach will be described in the next section for solving the above problem.

3  GENETIC ALGORITHM

S0: Chromosome representation

binary representation of length=$n$

(1: selected security, 0: non-selected security)

S1: Generate an initial populaton

Produce $pop_{size}$ chromosomes $v_k (k = 1, 2, \ldots, pop_{size})$ randomly.

S2: Repeat

S2-1: Evaluate fitness of each individual

fitness function is denoted as follows:

$$\alpha \cdot \frac{1}{r + 1} + (1 - \alpha) \cdot \frac{1}{(c - select)^2 + 1}.$$ 

Where,

$$r = a[i][j] \cdot (w[i] - \frac{1}{select} \cdot \text{chrom}[i]) \cdot (w[j] - \frac{1}{select} \cdot \text{chrom}[j]),$$
α: weights, c: numbers of '1' in the string.

S2-2: Selection (Roulette Selection method)
Let \( \{I_1, I_2, \ldots, I_N\} \) be set of population. Select solution \( I_i \) as a parent string according to selection probability \( P(I_i) \).

\[
P(I_i) = \frac{f(I_i)}{\sum_{j=1}^{N} f(I_j)}
\]

\( f(\cdot) \) is fitness value of the solution \( I \).

S2-3: Crossover
Two point crossover

S2-4: Mutation
Invert each bit in the solution with a small probability.

4 A NUMERICAL EXAMPLE

We test numerical examples in cases of \((n=14, m=7), (n=60, m=15)\) and real case \((n=225, m=30)\) by using NIKKEI 225 (1984–1988) for discussing the efficiency of proposed GA. In case of \((n=14, m=7)\), GA result is compared with known optimal solution.

5 CONCLUDING REMARKS

In this research, we developed an efficient GA to design an index fund with a given number of securities that minimizes the tracking error between the benchmark portfolio and the index fund. The proposed method can find the solutions as optimal or at least sub-optimal. Many problems are left for future research. In particular, it is necessary more theoretical analysis, simplification of operator and setting of parameter for better GA convergency.

参考文献


