Expert Systems with Applications 28 (2005) 371–379

Expert Systems with Applications

www.elsevier.com/locate/eswa

Using genetic algorithm to support portfolio optimization for index fund management

Kyong Joo Oh[a,*], Tae Yoon Kim[b], Sungky Min[a]

[a]Division of Business Administration, Hansung University, 389, 3-Ga, Samsun-Dong, Sungbuk-Gu, Seoul 136-792, South Korea
[b]Department of Statistics, Keimyung University, Daegu 704-701, South Korea

Abstract

Using genetic algorithm (GA), this study proposes a portfolio optimization scheme for index fund management. Index fund is one of popular strategies in portfolio management that aims at matching the performance of the benchmark index such as the S&P 500 in New York and the FTSE 100 in London as closely as possible. This strategy is taken by fund managers particularly when they are not sure about outperforming the market and adjust themselves to average performance. Recently, it is noticed that the performances of index funds are better than those of many other actively managed mutual funds [Elton, E., Gruber, G., & Blake, C. (1996). Survivorship bias and mutual fund performance. *Review of Financial Studies*, *9*, 1097–1120; Gruber, M. J. (1996). Another puzzle: the growth in actively managed mutual funds. *Journal of Finance*, *51*(3), 783–810; Malkiel, B. (1995). Returns from investing in equity mutual funds 1971 to 1991. *Journal of Finance*, *50*, 549–572]. The main objective of this paper is to report that index fund could improve its performance greatly with the proposed GA portfolio scheme, which will be demonstrated for index fund designed to track Korea Stock Price Index (KOSPI) 200.

Keywords: Index fund; Genetic algorithm; Portfolio

1. Introduction

Index funds are popular investment tools being used in modern portfolio management. Index funds are designed to mimic the behavior of the given benchmark market indices (e.g. the S&P 500 in New York, the FTSE 100 in London, the KOSPI 200 in Seoul, etc.). Thus, index funds are generally regarded as relatively stable and efficient investment tool compared with other mutual funds (Jensen, 1968; Sharpe, 1966).

The index fund strategy is based on the concept of the passive investment management. There are several interesting papers reporting the superior performance of the index funds compared with other actively managed portfolios (Elton, Gruber, & Blake, 1996; Gruber, 1996; Malkiel, 1995). In addition to the performances of index funds in terms of risk and return, index funds are also considered cost effective investment tool in the capital market (Hogan, 1994).

Index funds are composed of relatively small number of stocks. If we want to set up a perfect index fund, we need put every company included in the index into the index fund portfolio (e.g. 500 companies for S&P 500, 100 companies for FTSE 100, and 200 companies for KOPSI 200). However, it is costly and not practical to include every company in the index fund portfolio. Thus, index funds try to replicate the movement of the indices with a relatively small number of stocks.

This article proposes a genetic algorithm (GA) portfolio scheme for the index fund optimization. The scheme exploits GA and provides the optimal selection of stocks utilizing fundamental variables—standard error of portfolio beta given by formula (1), average trading amount, and average market capitalization. These fundamental variables are well-known core factors frequently used in analyzing and forecasting the stock market. Roughly speaking, the GA portfolio scheme consists of two steps. First, the stocks for the index fund are selected through working with the fundamental variables in each industry sector of the benchmark index. Second, the relative weights of the selected stocks are optimized through the GA process.

* Corresponding author. Tel.: +82 2 760 4423; fax: +82 2 760 4442.
E-mail address: johanoh@hansung.ac.kr (K.J. Oh).

0957-4174/$ - see front matter © 2004 Elsevier Ltd. All rights reserved.
doi:10.1016/j.eswa.2004.10.014

It will be shown that the portfolio scheme efficiently replicates the benchmark index with a relatively small number of stocks. Notice that the business of the efficient index fund management relies on the technique of replicating the benchmark index.

The proposed GA scheme is applied to Korea stock price index (KOSPI) 200 from Jan 1999 to Dec 2001. KOSPI 200 includes 200 major companies in 22 industry sectors, which are currently listed on the Korean Stock Exchange. The 200 companies cover general spectrum on the Korea Stock Exchange and KOSPI 200 is also the base index of KOSPI 200 futures contract, which is the most active futures contract on the Korea Stock Exchange.

This paper consists of five sections. Following this section, Section 2 provides a brief survey about portfolio theory, index funds, and GA. Section 3 discusses the detailed procedure of the proposed scheme and Section 4 reports empirical experiment results. Finally, Section 5 is devoted to the concluding remarks.

2. Literature review

2.1. Portfolio theory, index funds and index tracking

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, 1959), 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. 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 (Fama, 1970; Hakansson 1970, 1974; Merton, 1990; Mossin, 1969). After mean–variance portfolio theory, there was an enormous progress on portfolio theory and practice which include various practical applications introduced in portfolio formulation. Recently, it is found that low-cost passively managed index funds actually deliver the highest risk-adjusted returns in each category of mutual funds (Bogle, 1998).

Index fund management is a stock-allocation strategy equipped with index tracking skill which attempts to replicate the behavior of a given benchmark index. Index funds usually do not include every stock comprising the index. However, they are designed to copy the benchmark index with relatively small number of stocks, which can be easily managed and controlled in the capital market. Thus, the performance of the index fund critically depends on how well the index tracking skill replicates the benchmark index with only a subset of the stocks.

Index tracking, of course, involves tracking error (TE) which is measured by TE volatility, the sum of the deviations of returns of the replicating portfolio from the benchmark index. When fund managers formulate an index fund, they try to minimize the TE volatility level since it would produce as close as possible returns to the benchmark returns (Clarke, Krase, & Statman, 1994; Sharpe, 1971; Konno & Yamazaki, 1991). In general, there are several types of TE measures available, i.e. quadratic, linear and absolute, among which quadratic measure are preferred since it possesses a number of desirable statistical properties (Roll, 1992). Throughout this study, quadratic measure is employed.

2.2. Genetic algorithm

GA is a stochastic optimization technique invented by Holland (1975) and a search algorithm based on survival of the fittest among string structures (Goldberg, 1989). They applied the idea from biology research to guide the search to an (near-) optimal solution (Wong & Tan, 1994). The general idea was to maintain an artificial ecosystem, consisting of a population of chromosomes. In this study, each chromosome represents the weight of individual stock of portfolio and is optimized to reach a possible solution. Attached to each chromosome is a fitness value, which defines how good a solution the chromosome represents. By using mutation, crossover values, and natural selection, the population will converge to one containing only chromosomes with good fitness (Adeli & Hung, 1995). Recently, GA attracts much attention in portfolio formulations (Orito, Yamamoto, & Yamazaki, 2003; Xia, Liu, Wang, & Lai, 2000).

3. Scheme specification

3.1. Related variables

The proposed scheme is based on three fundamental variables: portfolio beta, trading amount, and market capitalization. These three variables are frequently used in portfolio management area, among which portfolio beta is especially the most important variable (Chang, 2004; Keim, 1999).

3.1.1. Portfolio beta

Let β_p be portfolio beta for a given portfolio p defined by

$$\beta_p = \frac{\mathrm{Cov}(r_p, I_m)}{\mathrm{Var}(I_m)}, \tag{1}$$

where r_p is the rate of return of the portfolio p and I_m is the rate of return of the benchmark index or the capital market m. It is well known that the portfolio beta measures portfolio volatility relative to the benchmark index or the capital market. Indeed, if a portfolio is well chosen such that returns of portfolio and benchmark index are highly correlated, then portfolio beta becomes the volatility ratio between the