A Universal Investment Strategy for Stock Markets Using Machine Learning

Seyed Mohammad Ziaemohseni*

Abstract

We introduce a sequential investment strategy for investing in stock markets. The only assumption we use is that the daily price relatives form a stationary and ergodic process. In this sense, our strategy is universal since we do not know and do not attempt to know precisely the exact statistics of the underlying market. We provide a few empirical results to show the performance of this strategy.

1 Introduction

Most investment strategies use information from the past behavior of the market to design a portfolio to be used for the next trading period. The goal is to maximize the wealth in the long run with the minimum knowledge of the underlying distributions of the stock prices. In this project, the only assumption we use is that the daily price relatives form a stationary and ergodic process. In this sense, our strategy is universal since we do not know and do not attempt to know precisely the exact statistics of the underlying market. It has been shown that there exist universal strategies that achieve the same asymptotic growth rate as if we had the full knowledge of the statistics of the underlying market.

2 Mathematical Setup

We consider a market of d assets and define the market vector $\mathbf{x}$ as:

$$\mathbf{x} = (x^{(1)}, x^{(2)}, ..., x^{(d)})^{\mathrm{T}} \in \mathbf{R}_{+}^{d}$$

$\mathbf{x}$ is vector of d nonnegative numbers representing the price relatives of the assets for a given period. More specifically, the j-th component $x^{(j)} \geq 0$ of $\mathbf{x}$ expresses the ratio of today's closing price to previous day's closing price for the asset j. Hence, $x^{(j)}$ is the factor by which capital invested in the j-th asset grows during the trading period.

*Information Systems Laboratory, Department of Electrical Engineering, Stanford University, Stanford, CA 94305, mohseni@stanford.edu

At the beginning of each period, the investor diversifies his capital according to a portfolio vector $\mathbf{b}$:

$$\mathbf{b} = (b^{(1)}, b^{(2)}, ..., b^{(d)})^{\mathrm{T}}$$

where the j-th component $b^{(j)}$ denotes the proportion of the capital being invested in asset j. During our analysis we assume that short selling is prohibited, this means that elements of $\mathbf{b}$ are nonnegative. We also assume a self-financing strategy which leads to the condition $\mathbf{1}^{\mathrm{T}}\mathbf{b} = 1$, where $\mathbf{1}$ is a vector of all 1's. With these notations, it is clear that if the investor's initial capital is S_0, then at the end of the first trading period his wealth is:

$$S_1 = S_0 \mathbf{b}^{\mathrm{T}} \mathbf{x}$$

We represent the evolution of market in time by a sequence of market vectors $\mathbf{x}_1, \mathbf{x}_2, ...$. To ease the notation, for $j \leq i$, we abbreviate by $\mathbf{x}_j^i$ the array of market vectors $(\mathbf{x}_j, ..., \mathbf{x}_i)$. Hence, at the beginning of the i-th trading period, our available history would be $\mathbf{x}_1^{i-1}$. The portfolio vector chosen by the investor at the beginning of the i-th trading period depends on the past behavior of the market. To emphasize this dependence, we denote this portfolio by $\mathbf{b}(\mathbf{x}_1^{i-1})$. Starting with an initial wealth S_0, after n trading periods the investor's wealth is:

$$S_n = S_0 \prod_{i=1}^{n} \mathbf{b}^{\mathrm{T}}(\mathbf{x}_1^{i-1})\mathbf{x}_i = S_0 e^{nW_n(\mathbf{B})}$$

where $W_n(\mathbf{B})$ denotes the average growth rate:

$$W_n(\mathbf{B}) = \frac{1}{n} \sum_{i=1}^{n} \log(\mathbf{b}^{\mathrm{T}}(\mathbf{x}_1^{i-1})\mathbf{x}_i)$$

For the first training period in which there is no market history available, we let $\mathbf{b}$ to be a uniform portfolio. The goal is to maximize S_n, or equivalently $W_n(\mathbf{B})$.

Two different approaches have been considered in the literature for maximizing $W_n(\mathbf{B})$. The first approach does not impose any stochastic model for the evolution of price relative vectors. Hence, the results hold for all possible sequences $\mathbf{x}_1^n$. The second approach assumes the market vectors are realizations of a random process and describes a statistical model for their evolution. In this case, one is always able to find an optimal investment strategy based on the distribution of the process which maximizes $W_n(\mathbf{B})$. However, describing a statistical model which accurately captures the behavior of market vectors has been proved to be an extremely difficult task.

In this project we adopt a compromise between these two approaches. We assume that the market sequence is a realization of a random process, but we do not assume any parametric structure on the distribution of this random process. In fact, the only statistical assumption we make is that the market process is stationary and ergodic. This allows us to deal with arbitrarily complex distributions, as long as they are stationary and ergodic.

Our proposed strategy allows a flexible way of extracting information from the history of the market. We denote the sequence of our portfolios by $\mathbf{B}^K = \{\mathbf{b}^K(.)\}$. $\mathbf{B}^K$ is constructed as follows. We define $c > 0$ to be a constant. Let $\mathbf{y}$ and $\mathbf{z}$ be two arbitrary vectors with arbitrary dimensions. We say $\mathbf{y}$ and $\mathbf{z}$ are similar to each other if and only if:

$$\|\mathbf{y} - \mathbf{z}\| \leq c \tag{1}$$

To determine our portfolio on the n-th trading period, we scan the available history of market vectors $\mathbf{x}_1^{n-1}$ and collect those vectors that followed a vector similar to $\mathbf{x}_{n-1}$ in a history list, denoted by $\mathbf{h}$. After scanning the whole history, we design our portfolio for the next trading day by solving the following optimization problem:

$$\mathbf{b}^K(\mathbf{x}_1^{n-1}) = \arg\max_{\mathbf{b} \in \triangle_d} \prod_{\mathbf{x}_i \in \mathbf{h}} \mathbf{b}^{\mathrm{T}} \mathbf{x}_i \tag{2}$$

where $\triangle_d$ is the probability simplex. If $\mathbf{h}$ is empty, we choose our portfolio to be a uniform portfolio.

REMARK It is important to choose a suitable value for c in (1) such that the number of nearest neighbors is neither too high nor too low. Also note that the optimization problem in (2) could be converted to a convex problem using a simple trick. Define $\mathbf{A} = diag(a_1, a_2, ..., a_n)$ to be a diagonal matrix, where n is the number of elements in $\mathbf{h}$. Now maximizing $\log det(\mathbf{A})$ subject to the set of constraints $\mathbf{b}^{\mathrm{T}} \mathbf{x}_i = a_i$ for $1 \leq i \leq n$ is an equivalent convex problem. In practice, instead of maximizing $\log det(\mathbf{A})$, we maximized $(det(\mathbf{A}))^{\frac{1}{n}}$ (which is still a convex problem) and used *cvx* to solve the problem.

3 Numerical Results

We tested this strategy on two different sets of financial data. In both cases the initial wealth S_0 is 1 dollar. The first data set includes daily prices of 10 stocks along a 22-year period (5651 trading days) ending in 1985. Figure 1 shows the wealth achieved by $\mathbf{B}^K$ during this period. We used the value $c = 0.050$ for our simulation. The final wealth is 5832 dollars, equivalent to a 47.3 percent annual interest rate.

The second set of data includes daily prices of 100 stocks along an 8-year period ending in December 2010. We used the value $c = 0.1826$. The final wealth is 36.1 dollars, equivalent to a 56.6 percent annual interest rate, as seen in Figure 2.

4 References

P. Algoet. Universal schemes for prediction, gambling, and portfolio selection. Ann. Prob., 20:901941, 1992.

L. Gyorfi, G. Lugosi, and F. Udina. Nonparametric kernel-based sequential investment strategies. Mathematical Finance, 16(2):337357, 2006.

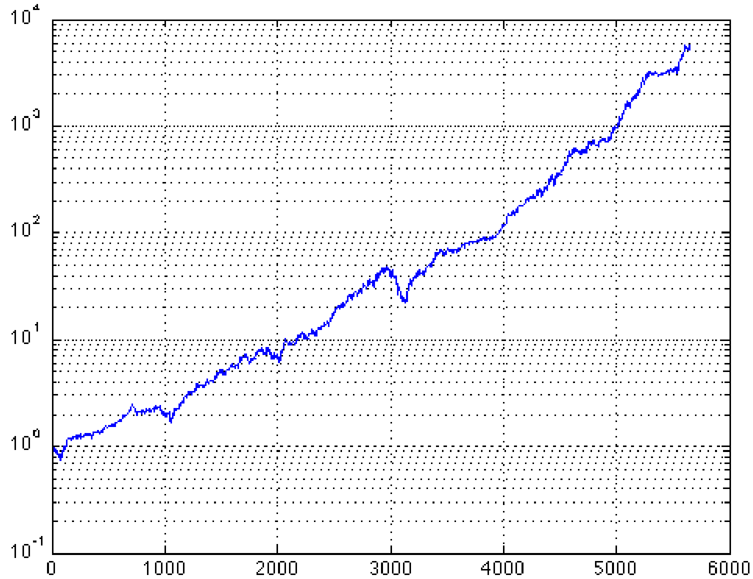

Figure 1: Plot of achieved wealth versus number of trading days for a basket of 10 stocks.

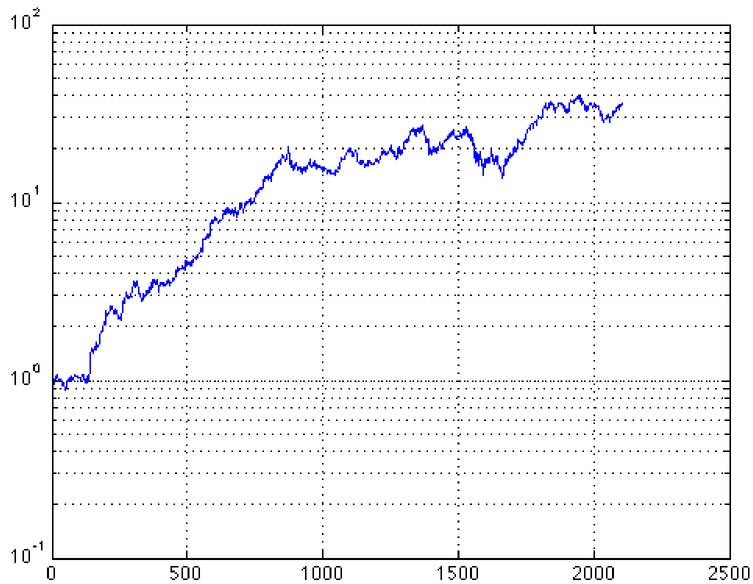

Figure 2: Plot of achieved wealth versus number of trading days for a basket of 100 stocks.