Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 22 (2013) 1172 – 1181

Procedia Computer Science

17th International Conference in Knowledge Based and Intelligent Information and Engineering Systems - KES2013

Predicting the security levels of stock investment by using the RMT-test

Xin Yang[a], Yuuta Mikamori[a], Mieko Tanaka-Yamawaki[a,] *

[a]Information and electronics Department, Graduate School of Engineering, Tottori University, Tottori 680-0945, Japan

Abstract

The authors propose to use the degree of randomness of high frequency price time series for the purpose of measuring the security levels of stock investments. The RMT-test is employed as a tool to measure the randomness. The data to be analyzed are the tick-wise price time series of selected stocks in the Tokyo Stock Exchange Market for three years from 2007 to 2009. The result shows that the stock of the highest randomness is a stable stock that belongs to the sector of electric/gas power supply, which turns out to be more profitable than the Nikkei Average Price throughout the following year. This indicates that the suitable stocks to invest under a bear market have higher randomness that belongs to the category of 'defensive' stocks, according to the new classification method introduced by Tanaka-Yamawaki, et. al., while the suitable stocks to invest under a bull market have lower randomness that belong to the category of 'outer demand' and 'market sensitive' stocks in the same classification method.

Selection and peer-review under responsibility of KES International

Keywords: Randomness, RMT-test, RMT-PCA, tick-wise stock price;

1. Introduction

By the rapid progress of digitization in recent years, acquisition and storage of huge amount of data has become available in various fields such as weather, health, finance and census. This situation requires development of effective technology, in order to analyze such heavy data for various purposes. The Principal Component Analysis based on the Random Matrix Theory (RMT-PCA) has been proposed as a technique to separate the random components and the correlated components, in order to extract useful information out of massive amount of data. The authors have proposed to measure the randomness of a given long time series called RMT-test [1], and have demonstrated the effectiveness of the RMT-test by measuring the randomness of

* Corresponding author. Tel.: +81-0857-31-5223; fax: +81-0857-31-0879.
E-mail address: mieko@ike.tottori-u.ac.jp.

Selection and peer-review under responsibility of KES International
doi:10.1016/j.procs.2013.09.204

the physical numbers and pseudo-random numbers [2]. However, the real advantage of the RMT-test is its applicability on the real-world data whose randomness is not very high. In this paper, we apply the RMT-test on tick-wise stock prices and attempt to use the randomness of price fluctuations as a new indicator for stock investments.

2. Review of RMT-test

The method used in this paper is to compare the eigenvalue distribution of the correlation matrix, between *N* pieces of length *L*, to the corresponding theoretical formula of the eigenvalue distribution derived from the random matrix theory in the limit of *N* and *L* going to infinity, keeping $Q = L/N$ as a constant. This method is applied to the stock market in 2002 by Plerou [3]. The outline of the method is as below.

A data sequence is cut into *N* pieces of equal length *L*, then shape them in an $L\times N$ matrix, by placing the first *L* elements in the 1st row of the matrix, and the next *L* elements in the 2nd row, and so on, by discarding the remainder if the length of the sequence is not divisible by *L*. Then each column of the matrix is normalized to have zero mean and single variance. By multiplying this matrix with its own transverse matrix, the correlation matrix *C* is constructed, which is a symmetric $N\times N$ matrix whose (i, j) element is the inner products between the *i*-th and the *j*-th columns of the $L\times N$ matrix. All of the *N* eigenvalues of the correlation matrix *C* is obtained by numerical calculation. The randomness of the sequence is measured by comparing the eigenvalue distribution to the corresponding theoretical formula, called Marcenko-Pastur distribution:

$$P_{RMT}(\lambda) = \frac{Q}{2\pi}\frac{\sqrt{(\lambda_+ - \lambda)(\lambda - \lambda_-)}}{\lambda} \quad (1)$$

$$\lambda_\pm = (1 \pm Q^{-1/2})^2 \quad (2)$$

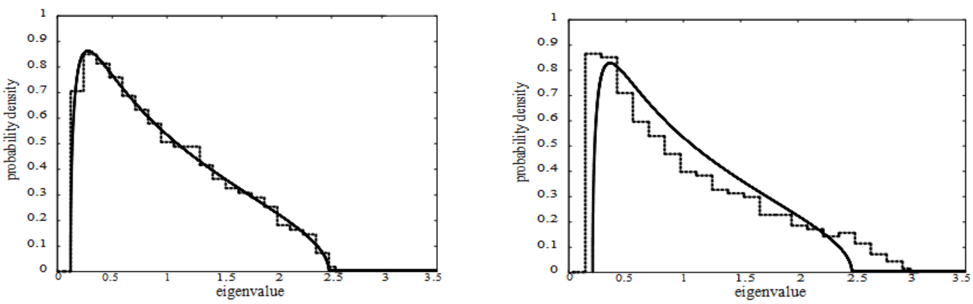

Fig.1. Example of the Qualitative evaluation of RMT-test (left: passed, right: failed)

The quantitative evaluation based on the moment method [4] compares the k-th moment of the obtained eigenvalues

$$m_k = \frac{1}{N}\sum_{i=1}^{N}\lambda_i^k \quad (3)$$

with the corresponding theoretical formula obtained from P_{RMT}

$$\mu_k = E(\lambda^k) = \int_{\lambda_-}^{\lambda_+}\lambda^k P_{RMT}(\lambda)d\lambda \quad (4)$$

The difference between m_k and μ_k represents the degree of randomness of the data sequence. The authors