Physica A 453 (2016) 150–161

Contents lists available at ScienceDirect

Physica A

journal homepage: www.elsevier.com/locate/physa

How long the singular value decomposed entropy predicts the stock market? — Evidence from the Dow Jones Industrial Average Index

Rongbao Gu [a,*], Yanmin Shao [b]

[a] School of Finance, Nanjing University of Finance & Economics, Nanjing 210046, PR China

[b] Center for Forecasting Science, Chinese Academy of Sciences, Beijing 100190, PR China

HIGHLIGHTS

- A new concept of multi-scales singular value decomposition entropy is proposed.
- The predictive power of the entropies to DJIAI is analyzed for different time scales.
- The entropy has short-term (not more than one month) predictive power for DJIAI.

ARTICLE INFO

Article history:
Received 6 November 2015
Received in revised form 8 January 2016
Available online 18 February 2016

Keywords:
Stock market
Prediction
DCCA
Entropy

ABSTRACT

In this paper, a new concept of multi-scales singular value decomposition entropy based on DCCA cross correlation analysis is proposed and its predictive power for the Dow Jones Industrial Average Index is studied. Using Granger causality analysis with different time scales, it is found that, the singular value decomposition entropy has predictive power for the Dow Jones Industrial Average Index for period less than one month, but not for more than one month. This shows how long the singular value decomposition entropy predicts the stock market that extends Caraiani's result obtained in Caraiani (2014). On the other hand, the result also shows an essential characteristic of stock market as a chaotic dynamic system.

1. Introduction

The efficient market hypothesis proposed by Fama [1] in 1965 is an important foundation of the modern finance theory. A market is called to be week-efficient market if the price of stock reflects all past publicly available information. If a market is week-efficient, then the price of stock should walk randomly and the stock market cannot be predicted. In fact, there is neither absolutely efficient market nor absolutely inefficient market in the real market. Any stock market is on a condition between efficient and inefficient market, namely, all the stock markets are complex dynamic systems [2]. Therefore, we should understand stock market from the view of complex science. With the development of nonlinear science, a lot of nonlinear scientific ideas, methods and technologies have been applied into the research of stock market. A detailed review is presented in Ref. [3].

Recently, Caraiani [4] did a research about the predictive power of the singular value decomposition entropy for stock market. He found that the singular value decomposition entropy has predictive power for the Dow Jones Industrial Average

* Corresponding author.
E-mail address: rbgu@sina.com (R. Gu).

http://dx.doi.org/10.1016/j.physa.2016.02.030

Index. Following Caraiani, Gu et al. [3] investigated predictive power of the singular value decomposition entropy for Chinese Shenzhen stock market. They found that, the predictive power of the entropy for the Shenzhen component index is affected by the structural break in stock market, and the entropy only has predictive power for the Shenzhen component index after the reform of non-tradable shares.

Entropy is a kind of measurement of the uncertainty of system state as well as an index reflecting the complex degree of signals. The larger the value of entropy is, the more complex the signals are, and the more abundant information the signals contain. However, the entropy proposed by Caraiani [4] is based on the singular value decomposition for the Pearson correlation coefficient matrix. It is known that the Pearson correlation coefficients only describe linear correlation of the series. So, the singular value decomposition entropy based on the matrix composed of Pearson correlation coefficients only reflects simple (linear) correlation information between the prices of stocks. In fact, the correlations between the prices of stocks are quite complex as the stock market is a quite complex dynamic system. These complex correlations between the prices of stocks usually display nonlinear relations that cannot be described by the singular value decomposition entropy based on the Pearson correlation coefficients matrix. So, one cannot better understand the predictive power of the singular value decomposition entropy for stock index employing the notion proposed by Caraiani [4].

Using the technique of multifractal analysis, Zebende [5] proposed a new nonlinear correlation coefficient, *DCCA cross-correlation coefficient*, which can quantify the level of cross-correlation between two series with non-stationary and non-Gaussian distribution. The outstanding advantage of this coefficient is that it can measure the correlations between two financial time series with the different time scales. Based on the DCCA cross-correlation coefficient, we introduce in this paper a new concept of multi-scales singular value decomposition entropy, and investigate its predictive power for the Dow Jones Industrial Average Index employing linear and nonlinear Granger causality test with different time scales. This can help us to better understand the predictive power of the singular value decomposition entropy for stock index with different time scales.

This paper is organized as follows. The following section is dedicated to an introduction about the DCCA cross-correlation coefficient and multi-scales singular value decomposition entropy. Section 3 introduces the linear and nonlinear Granger causality test used to analyze the predictive power of entropy for stock market. Data and its basic statistical analysis are presented in Section 4. The analysis about predictive power of multi-scales singular value decomposition entropy for the Dow Jones Industrial Average Index with different time scales is presented in Section 5. The last section is a brief conclusion.

2. Correlation matrix and singular value decomposition entropy

2.1. Pearson correlation matrix

The most popular measure of correlation of two time series is Pearson correlation coefficient, which is the ratio of their co-variance to the product of the two standard deviations.

Assume that S_k is kth component stock in the index of some stock market and S_{kt} is the closing prices of S_k at term t. The Pearson correlation coefficient between stocks S_i and S_j is calculated as follows:

$$P_{i,j} = \frac{\langle (S_{it} - \langle S_{it} \rangle)(S_{jt} - \langle S_{jt} \rangle) \rangle}{\sigma_i \sigma_j} \tag{1}$$

where $\langle \cdot \rangle$ denotes the mean value and σ_k denotes the standard deviation of the stock S_k.

The Pearson correlation coefficient $P_{i,j}$ describes the degree of linear correlation between prices of stock S_i and S_j. The bigger the absolute value of Pearson correlation coefficient is, the stronger the linear correlation of the two stock prices could be.

We denote by $\boldsymbol{P}$ the matrix composed of Pearson correlation coefficients $P_{i,j}$, i.e.,

$$\boldsymbol{P} = (P_{i,j}), \tag{2}$$

which is called correlation matrix of index of the stock market. The correlation matrix offers to us the information about the degree of linear correlation between component stocks of index of the stock market.

The application of correlation matrix can retrospect to Mantegna's work [6]. He proposed an information filtering method using minimum spanning tree. This work has been generalized by Tumminello et al. [7], who proposed an information filtering method using maximal planar graph and suggested that the maximal planar graph contains more information than minimum spanning tree. As a result of huge effect of financial crisis to global economy, some authors tried to study how to detect the stock dynamic change and probability of future crisis through information from correlation matrix (Song et al. [8], Kenett et al. [9] and Peron et al. [10]).

2.2. DCCA cross-correlation matrix

The Pearson correlation coefficient is not robust. It is affected by the outliers, also non-stationarity or non-Gaussian distribution. Recently, Zebende [5] proposed a new nonlinear correlation coefficient, *the DCCA cross-correlation coefficient*, in order to quantify the level of cross-correlation between two series with non-stationary and non-Gaussian distribution.