Commun Nonlinear Sci Numer Simulat 52 (2017) 77–90

Contents lists available at ScienceDirect

Commun Nonlinear Sci Numer Simulat

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

Research paper

Multivariate multiscale entropy of financial markets

Yunfan Lu*, Jun Wang

School of Science, Beijing Jiaotong University, Beijing 100044, PR China

ARTICLE INFO

Article history:
Received 31 December 2016
Revised 10 April 2017
Accepted 24 April 2017
Available online 25 April 2017

Keywords:
Multivariate multiscale entropy analysis
Nonlinear complexity behavior
Globe stock market
Multivariate volatility series
Shuffled multivariate return

ABSTRACT

In current process of quantifying the dynamical properties of the complex phenomena in financial market system, the multivariate financial time series are widely concerned. In this work, considering the shortcomings and limitations of univariate multiscale entropy in analyzing the multivariate time series, the multivariate multiscale sample entropy (MMSE), which can evaluate the complexity in multiple data channels over different timescales, is applied to quantify the complexity of financial markets. Its effectiveness and advantages have been detected with numerical simulations with two well-known synthetic noise signals. For the first time, the complexity of four generated trivariate return series for each stock trading hour in China stock markets is quantified thanks to the interdisciplinary application of this method. We find that the complexity of trivariate return series in each hour show a significant decreasing trend with the stock trading time progressing. Further, the shuffled multivariate return series and the absolute multivariate return series are also analyzed. As another new attempt, quantifying the complexity of global stock markets (Asia, Europe and America) is carried out by analyzing the multivariate returns from them. Finally we utilize the multivariate multiscale entropy to assess the relative complexity of normalized multivariate return volatility series with different degrees.

1. Introduction

Nonlinear science, an important and challenging subject, has got many contributions about the modeling and the behavior analysis in many branch field, such as the financial market analysis [1–3], the physiologic system [4], biological signal [5], laster doppler flowmetry system [6], oil-water flow [7], the multi-hydro-turbine governing systems [8–10] and human gait dynamics [11]. It is also noteworthy that the empirical studies which focus on the discovery of nonlinear features in financial dynamics have been attracting extensive attentions in recent years. Considered as a complex evolved nonlinear system, the financial market has the empirically observed stylized facts, such as fat tails phenomenon and power law of logarithmic returns [1,12–14], long-term memory and volatility clustering [15,16], autocorrelation and cross-correlation [17,18], multifractality of volatility and multifractal behaviors [2,19,20], complexity of financial systems [3,18]. These studies, aiming at extending and consolidating the known stylized facts, are also essential for risk management and has aroused wide attention from the community of researchers in nonlinear financial dynamical systems. Nowadays, many research literatures have arisen to explore the nonlinear natures of financial markets [2,14,19,21–25]. One of the most significant aspects is to quantify the complexity of financial time series. The recently introduced multiscale entropy (MSE) approach proposed by Costa et al. [4] has become a useful and popular method to quantify the complexity of signals in different

* Corresponding author.
E-mail address: yunfanlu@bjtu.edu.cn (Y. Lu).

http://dx.doi.org/10.1016/j.cnsns.2017.04.028
1007-5704/

research fields, such as the biomedical and physiologic time series [5,6], the vibration of rotary machine [26], the human gait pattern at different ages [11,27], and the financial time series [3]. The MSE method (that evaluates sample entropy from univariate time series, coarse grained on multiple scales, and thus indicates long-range correlations over a range of temporal scales within a complex system) has been proven to be able to distinguish physiological time series with different degrees of complexity. However, it is also shown that this MSE method has some shortcomings and limitations, especially it is not suited for multivariate time series that are routinely recorded in experimental, physical and biological systems thanks to the recent developments in sensor technology. In the MSE method, multivariate time series are considered as a set of individual time series, and they are separately calculated in each variable to assess their complexity. However, this is only adequate if each variable in a multivariate signal is statistically independent or at the very least uncorrelated (which is usually not the case). To generalize the univariate MSE to the multivariate case, Ahmed and Mandic introduced multivariate sample entropy (MSampEn), and evaluate its evolution over multiple time scales to perform the multivariate multiscale entropy (MMSE) analysis [28]. The introduced multivariate multiscale entropy (MMSE) has been shown substantial advantages in operating on any number of data channels simultaneously, and providing a dynamical complexity measure for multivariate data observed from the same system, especially if there is a large degree of uncertainty or coupling underlying the measured system dynamics. The approach has been supported by simulations on both synthetic and real world multivariate processes, such as gait, wind, and physiological data [7,28,29]. However, the multivariate multiscale entropy (MMSE) has not been applied in analyzing the multivariate time series in financial market.

In the present work, we detect the effectiveness and advantages of the MMSE method in complexity analysis of multivariate time series and try to determine appropriate parameters. For example, the effective length of the multivariate time series with small experimental error on two well-known synthetic noise signals, white noise (with uncorrelated fluctuation) and $1/f$ noise (with correlated fluctuations), is studied. Then the MMSE method is applied in the complexity analysis of China stock markets with four generated trivariate return series for each stock trading hour, also with their shuffled multivariate return series and absolute multivariate return series. Finally, a new approach of analyzing the complexity of stock markets for three major regions (Asia, Europe and America) of the world is carried out by studying the multivariate stock returns (which are derived from the logarithmic returns of the stock indices) from them. The multivariate return series with different degrees of volatility of three regions are also quantified by using the MMSE analysis. The logarithmic returns of a stock index is defined as:

$$R(t) = \ln \mathscr{P}(t) - \ln \mathscr{P}(t - \Delta t) \tag{1}$$

where $\mathscr{P}(t)$ denote the price of a given stock market index, and the corresponding price return $R(t)$ is defined as the change of the logarithm of the price in a given time interval Δt [30–32].

2. Multivariate multiscale entropy analysis

2.1. Multivariate sample entropy

Recently, Ahmed & Mandic [28] proposed the multivariate sample entropy (MSampEn), which enables entropy calculation for multichannel data by taking into account both within- and cross-channel dependencies, and then introduced it into the multivariate multiscale entropy (MMSE) analysis. We show the multivariate sample entropy (MSampEn) in Algorithm 1, and by extending the standard univariate sample entropy in Ref. [4], we introduce the multivariate multiscale entropy (MMSE) analysis in Algorithm 2. In Algorithm 1, the multivariate sample entropy method is based on the estimation of the conditional probability that two similar sequences will remain similar when the next data point is included. This is achieved by calculating the average number of neighboring delay vectors for a given tolerance level r and repeating the process after increasing the embedding dimension from m to $m+1$. Fig. 1 illustrates the principle behind multivariate sample entropy calculation with two high-frequency (five-minute interval) returns of Shanghai Stock Exchange (SSE) and Shenzhen Stock Exchange (SZSE) from February 16, 2015 to February 18, 2015, separately denoted as $x(t)$ and $y(t)$. In order to show clearly, each of the return series only shows 100 points in Fig. 1(a). For the illustration of principle, when the embedding dimension $m = 2$, we assume that the time lag vector $\tau = [1, 1]$ and the embedding vector $\mathbf{M} = [1, 1]$, then the composite bivariate delay vectors are $[x(t), y(t)]$ as shown in Fig. 1(b), where t denotes the time as sample index. In the process of MSampEn calculation, for any such vector (e.g., $[x(42), y(42)]$), we need to count the number of neighbors which are within a distance r (tolerance level), illustrated by a circle centered at $[x(42), y(42)]$ with radius r in Fig. 1(b). And then, upon increasing the embedding dimension from $m = 2$ to $m = 3$, the embedding vector $\mathbf{M}$ evolves to two new embedding vectors, $\mathbf{M} = [2, 1]$ and $\mathbf{M} = [1, 2]$, according to Step 4 in Algorithm 1. That is the reason that we have two possible subspaces: (i) the subspace of all the vectors $[x(t), x(t+1), y(t)]$ is shown in Fig. 1(c), and (ii) the subspace of all the vectors $[x(t), y(t), y(t+1)]$ is shown in Fig. 1(d). Similarly, for any such vector (e.g., $[x(42), x(43), y(42)]$), we also need to count the number of neighbors which are within a distance r (tolerance level), illustrated by a sphere centered at $[x(42), x(43), y(42)]$ with radius r in Fig. 1(c). For any such vector (e.g., $[x(42), y(42), y(43)]$) in Fig. 1(d), we should do the similar counting. We employ this rigorous approach to compare composite delay vectors (to find the neighbors) not only within each subspace but also across all the subspaces, thus fully catering for both within- and cross-channel correlations.