Physica A 501 (2018) 264–275

Contents lists available at ScienceDirect

Physica A

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

Extended AIC model based on high order moments and its application in the financial market

Xuegeng Mao, Pengjian Shang *

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

HIGHLIGHTS

- The Akaike Information Criteria(AIC) is extended to detect the impacts of skewness and kurtosis.
- The extended AIC algorithm is combined with multiscale wavelet analysis to reduce the influence of noise.
- AIC plane is created to analyze the relationships among three kinds of calculation of AIC.
- This technique is tested to analyze the singularities and classification of stock indices.

ARTICLE INFO

Article history:
Received 13 November 2017
Received in revised form 7 January 2018
Available online 15 March 2018

Keywords:
Akaike Information Criteria (AIC)
Skewness
Kurtosis
Wavelet transform
Financial market

ABSTRACT

In this paper, an extended method of traditional Akaike Information Criteria(AIC) is proposed to detect the volatility of time series by combining it with higher order moments, such as skewness and kurtosis. Since measures considering higher order moments are powerful in many aspects, the properties of asymmetry and flatness can be observed. Furthermore, in order to reduce the effect of noise and other incoherent features, we combine the extended AIC algorithm with multiscale wavelet analysis, in which the newly extended AIC algorithm is applied to wavelet coefficients at several scales and the time series are reconstructed by wavelet transform. After that, we create AIC planes to derive the relationship among AIC values using variance, skewness and kurtosis respectively. When we test this technique on the financial market, the aim is to analyze the trend and volatility of the closing price of stock indices and classify them. And we also adapt multiscale analysis to measure complexity of time series over a range of scales. Empirical results show that the singularity of time series in stock market can be detected via extended AIC algorithm.

1. Introduction

It is well-known that financial markets have been considered as complex dynamic systems consisting of complex factors [1–3]. Stock market indices are important measures of financial markets and are normally used to benchmark the performance of stock portfolios. And a number of statistical mechanics have been used to observe the features of stock markets, such as correlation function [4], multifractal [5], graph-theoretical approach [6] and Random Matrix Theory [7]. It is significant to investigate the knowledge of financial areas. Recently, segmentation studies on such complex systems have attracted much attention, and many great numerical results have been obtained.

Segmentation algorithm is an important distinction because the time series is inherently dynamic [8]. Stanley initially proposed the time-series segmentation as a model and a method, in whose article the segments were regarded as falling into

* Corresponding author.
E-mail addresses: 16118426@bjtu.edu.cn (X. Mao), pjshang@bjtu.edu.cn (P. Shang).

https://doi.org/10.1016/j.physa.2018.02.159
0378-4371/

classes in 1982 [9]. The segmentation algorithm was obtained by utilizing the relaxation method to maximize the resulting likelihood function. Later lots of articles were published to extend and complement the theory. Pedro et al. presented the compositional segmentation based on the Jensen–Shannon entropic divergence [10]. The method was conceptually simple and efficient, which could decompose sequences into homogeneous patches and was successfully used to explore the correlations of DNA sequences. Then the online segmentation of time series was introduced by applying least-squares approximation and orthogonal polynomials, in which method the results can obtain high accuracy [11]. For more excellent review of time series segmentation, see reference [12–16].

Recently, model evaluation criteria have attracted more and more attention in statistical paper [17–20]. It is necessary to solve the problem that how certain we can choose the best approximating model by a proper model evaluation criterion. The non-normalized Kullback–Leibler information is a useful method to separate the models [21]. However, Kullback–Leibler information cannot be directly observed or estimated since it relies on the true distribution and the model parameters which consequently are unknown. Later, Akaike information criteria (AIC) was developed by Akaike to numerically express the Kullback–Leibler discrepancy between the model generating the data and the approximating model, which is an asymptotically estimator and ranks competing model by representing the amount of information lost [22–25]. And the AIC could be considered as a kind of segmentation of time series. Later, Maeda proposed a different AIC algorithm, which could be directly calculated according to the time series. The method simplified the process of calculation significantly and did not need to fit the models with the AR processes [26].

Since the AIC for model selection is increasingly used in wildlife literature, such as ecology [27], PD location [28] and P-wave arrival detection [29], it is rarely applied in the financial market. With some modifications, the algorithm can be utilized to stock markets. Moreover, it is also interesting to explore the properties between high order statistical moments and AIC method according to the equation proposed by Maeda [26]. For various kind of reasons, the second-order moment, such as variance, is often used to deal with some statistical problems. As we all know, the second-order moment is exhaustively describing the Gaussian stochastic processes. Nevertheless, it may cause the loss of information. The necessity of investigating the moment of higher orders than the variance cannot be ignored. In recent years, time series analysis is gradually inclined to use high order moments, like skewness (normalized third-order central moment) and kurtosis (normalized fourth-order central moment) [30]. These measures are powerful in many aspects, for example, choosing a member of a family such as from the Karl Pearson family, developing a test of normality, and exploring the robustness of the standard normal theory procedures [31]. However, the presented literatures indicate that there is little research on higher moments than the fourth due to the difficulty of coping with the higher moments [32].

Skewness describes the degree of asymmetry of a variable distribution around its mean. The value of skewness will be around zero if the distribution is symmetric. A distribution with an asymmetric tail that is longer on the left will obtain a negative skewness. A distribution with an asymmetric tail that is longer on the right will obtain a positive skewness. Kurtosis is a measure of the relative peak or flatness of a distribution, compared with the normal distribution of the same variance. The kurtosis is sometimes defined the normalized fourth central moment minus 3. The value of kurtosis will be high when the distribution has heavy tails while the value will be low when the distribution has light tails.

For the standard AIC method, it is supposed that a signal can be divided into locally stationary segments, and that the intervals before and after the onset time are two distinct stationary processes [33]. However, concerning most of the time series, the existence of noise is not strange and the AIC algorithm may perform bad [28]. To avoid the problem as much as possible, we introduce the application of the wavelet transform to the guide of work of the AIC algorithm. Wavelet analysis is a relatively new statistical tool in data processing and decomposes a given signal into different frequency orthogonal components, after which each component is coped with a resolution matched to its scale, making it useful in distinguishing seasonality and detecting local and global dynamic properties of a signal [34]. And the wavelet analysis is not limited by the stationary assumption.

The aim of this paper is to expand the studies by Akaike. The structure of the paper is as follows. Section 2 first illustrates the concept of AIC, then presents the algorithm of the extended AIC based on high order moments. In Section 3 we show the data description. Section 4 demonstrates the application of financial time series. Finally, Section 5 summarizes the conclusions.

2. Methodology

2.1. Akaike information criteria (AIC)

AIC is a robust measure to select the best approximating model of given models and is an autoregressive (AR) time picking algorithm. This criterion has been widely used in seismology to investigate the onset of seismic waves [33]. There are two different formulas when calculating AIC. One is

$$AIC = -2ln(L) + 2k \tag{2.1}$$

where L is the maximum likelihood estimate for the model and k is the number of fitted parameters in the model. The other is

$$AIC = nln(\frac{RSS}{n}) + 2k \tag{2.2}$$