Physica A 392 (2013) 2460–2469

Contents lists available at SciVerse ScienceDirect

Physica A

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

Conditional independence graph for nonlinear time series and its application to international financial markets

Wei Gao [a,*], Hongxia Zhao [b]

[a] School of Statistics, Xi'an University of Finance and Economics, Xi'an Shaanxi 710061, China

[b] School of College English, Xi'an University of Finance and Economics, Xi'an Shaanxi 710061, China

ARTICLE INFO

Article history:
Available online 14 March 2013

Keywords:
Nonlinear time series
Conditional independence graphs
Conditional mutual information
Financial markets

ABSTRACT

Conditional independence graphs are proposed for describing the dependence structure of multivariate nonlinear time series, which extend the graphical modeling approach based on partial correlation. The vertexes represent the components of a multivariate time series and edges denote direct dependence between corresponding series. The conditional independence relations between component series are tested efficiently and consistently using conditional mutual information statistics and a bootstrap procedure. Furthermore, a method combining information theory with surrogate data is applied to test the linearity of the conditional dependence. The efficiency of the methods is approved through simulation time series with different linear and nonlinear dependence relations. Finally, we show how the method can be applied to international financial markets to investigate the nonlinear independence structure.

1. Introduction

In recent years increasing research studies involved the interaction structure between national stock markets and many empirical works in financial data used graphical models, e.g. Refs. [1,2]. Allali et al. [3] applied a partial correlation graph to study the interaction structure between international markets and investigated the conditional interaction structure between the most important international financial returns. They suggested using graphical interaction models to analyze the partial associations of stock markets. However, nonlinearity is a typical feature of financial data. Thus graphical models which can present nonlinear relations should be applied to analyze financial time series.

Graphical models have become an important tool for analyzing multivariate data. Through merging the probabilistic concept of conditional independence with graph theory by representing possible dependence among the variables of a multivariate distribution in a graph. Graphical models led to simple graphical criteria for identifying and visualizing the conditional independence relations that are implied by a model associated with a given graph. For an introduction to graphical models we refer to the monographs by Whittaker [4], Edwards [5], and Cox and Wermuth [6]; a mathematically more rigorous treatment can be found in Ref. [7]. From the study in Refs. [8,9] there has been an increasing interest in the use of graphical modeling techniques to analyze multivariate time series (e.g., Refs. [10–14]). However, most of these works have been restricted to the analysis of linear interdependence among the variables whereas the recent trend in time series analysis has shifted towards non-linear parametric and non-parametric models (e.g., Ref. [15]). Eichler [11] introduced a graphical time series model for the analysis of dynamic relationships among variables in multivariate time series. The modeling approach is based on the notion of strong Granger causality and can be applied to time series with non-linear

* Corresponding author.
E-mail address: gaoweiufe@yahoo.cn (W. Gao).

0378-4371/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
doi:10.1016/j.physa.2012.07.002

dependence. Chu and Glymour [16] proposed an approach to learn a class of additive nonlinear time series based on additive model regression. Gao and Tian [17] presented statistics based on conditional mutual information to test the conditional Granger causality and instantaneous dependence between the components of multivariate time series.

In this paper, we present conditional independence graphs for visualizing the direct dependence between multivariate time series. The vertexes represent the components of a multivariate time series and edges denote direct dependence (both linear and nonlinear) between corresponding series. Granger and Lin [18] and Diks and Manzan [19] applied information theoretic quantities to test the lag dependence in univariate time series. We use a statistic based on conditional mutual information to test the direct conditional independence between the component series. The statistic is estimated by correlation integral and the significance of the test statistics is determined by a bootstrap procedure.

In addition, a statistic based on general and linear conditional mutual information is proposed to test the nonlinearity of the relations between the time series. A bootstrap method combined with surrogate data is used to determine the significance of the test statistics. The bootstrap time series preserving all the linear properties of the original series is generated by means of surrogate data rather than an estimated VAR model, which avoids the calculating complexity and errors of estimating the VAR model.

This paper is developed as follows. Section 2.1 defines the conditional independence graph for multivariate nonlinear time series. In Section 2.2, the statistic to test conditional independence of the components is estimated by a nonparametric method based on correlation integrals and the steps for constructing a graph model from observed series are listed. Section 3.1 introduces the method of Palus(1996) to test nonlinearity in a system. In Section 3.2, the information theoretic statistic and the bootstrap procedure with surrogate data are presented to test the nonlinearity of the conditional dependence detected in Section 2.2. The effectiveness of the methods is shown by simulation examples with linear and nonlinear relations in Section 4. Section 5 provides the results of the methodology analyzing the interdependence structure in financial markets. Finally some concluding remarks are provided in Section 6.

2. Conditional independence graph for multivariate time series

2.1. Conditional independence graph

Suppose $X(t) = (X_1(t), \ldots, X_k(t))^T$, $t \in \mathbb{Z}$ is a k-dimensional time series. A graph $G = (V, E)$ for $X(t)$ consists of two sets, one set of vertices $V = \{1, \ldots, k\}$ and a set of edges $E \subset \{(a, b) \in V \times V\}$. We only consider undirected graphs, i.e. we assume $(a, b) \in E$ whenever $(b, a) \in E$. A visualization of a undirected graph can be accomplished by drawing a circle for each vertex and connecting each pair a, b for which $(a, b) \in E$ by a line. This line illustrates undirected edges and stands for symmetrical interaction. Indirect associations can result from subsequent direct influences. Two variables a and b are conditionally independent given all other variables if they are not connected via an edge, i.e. $(b, a), (a, b) \in E$. This is the pairwise Markov property for undirected graphical models. If two variables a and b are connected by a path, that is if vertices $a = a_0, \ldots, a_l = b$ exist such that there is an edge between each pair of successive vertices, there is some relationship between them, possibly mediated by other variables. A connectivity component of an undirected graph is a maximal subset of pairwise connected variables. The global Markov property says that two sets of variables A and B are conditionally independent given a set of variables $C \subset V$ if C separates A and B in, i.e. if any path between variables $a \in A$ and $b \in B$ necessarily contains at least one variable $c \in C$.

Without loss of generality, we consider the conditional independence of series X_a and X_b for an example. In order to discriminate between a direct and induced relation between $X_a(t)$ and $X_b(t)$ by analogy to graphical models for multivariate data, we need to subtract the effects of the remaining components of $X(t)$ from $X_a(t)$ and $X_b(t)$. Let $Y_{ab}(t) = (X_j(t), j \neq a, b)$, $\mathscr{X}_a = (X_a(t), t \in \mathbb{Z})$ and $\mathscr{Y}_{ab} = (Y_{ab}(t), t \in \mathbb{Z})$.

Definition 1. Let $X(t) = (X_1(t), \ldots, X_k(t))^T$, $t \in \mathbb{Z}$ be a multivariate stationary time series. Graph $G = (V, E)$, where $V = \{1, \ldots, k\}$ is the corresponding set of vertices with i denotes the component series $X_i(t)$. $(a, b) \notin E$ if and only if the relation defined as followed is satisfied,

$$\mathscr{X}_a \perp\!\!\!\perp \mathscr{X}_b | \mathscr{Y}_{ab} :\Leftrightarrow X_a(t) \perp\!\!\!\perp X_b(t+u) | \mathscr{F}_{ab} \quad (1)$$

where $\mathscr{F}_{ab}$ denotes the σ-algebra generated by $\mathscr{Y}_{ab}$. Then $G = (V, E)$ is a conditional independent graph for multivariate time series $X(t)$.

A graphical model building by Definition 1 satisfies the pairwise Markov property. Dahlhaus [9] and Eichler [11] proved that under some assumption the pairwise Markov property is equivalent to the global Markov property for linear and nonlinear time series respectively. Throughout the paper, we assume that the series $X(t)$ satisfies the following conditions and then the two Markov properties are equivalent.

(A1) $X(t) = (X_1(t), \ldots, X_k(t))^T$, $t \in \mathbb{Z}$ is a stationary, mixing stochastic process on some probability space $(\Omega, \mathscr{F}, \mathbb{P})$.

(A2) For $\forall A \subset V$, the conditional distribution $\mathbb{P}(X_A(t)|X_{V\setminus A}(t+u))$ has a regular version that is almost surely absolutely continuous with respect to some product measure υ on $\mathbb{R}^d$ with υ-a.e. positive density.