Physica A 486 (2017) 701–710

Contents lists available at ScienceDirect

Physica A

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

Directed information graphs for the Granger causality of multivariate time series

Wei Gao *, Wanqi Cui, Wenna Ye

School of Statistics, Xian University of Finance and Economics, China

HIGHLIGHTS

- A definition of Granger causality is proposed.
- A decomposition of conditional directed information is proved.
- The directed information graphs are presented to describe the Granger causality.
- The method for structure learning of the graph models is investigated.

ARTICLE INFO

Article history:
Received 27 March 2017
Available online 10 June 2017

Keywords:
Directed information
Granger causality
Multivariate time series
Graphical model

ABSTRACT

In this paper, we investigate the links between (strong) Granger causality and directed information theory for multivariate time series. Based on the decomposition of conditional directed information, we propose a definition of Granger causality including instantaneous variables in the conditional set, which can avoid the spurious causality. The directed information graphs are presented to describe the Granger causality and instantaneous coupling. The structure learning of the graph models is based on the Leonenko's k-nn estimator of the statistics and a permutation test of the significant. Finally, we demonstrate the numerical implementation of these techniques on linear and nonlinear time series.

1. Introduction

Granger causality was developed in 1960s [1] and has been widely used in economics, biology, neuroscience, etc. The fundamental idea of Granger's approach is that the past and the present may cause the future, but the future cannot cause the past [2]. Recently, the graphical model has successfully provided a general framework for modeling conditional independence relations in multivariate time series. Dahlhaus [3] and Eichler [4] proposed the use of graphical modeling to describe Granger causality. Eichler [4] proposed a precise definition of Granger causality graphs and investigated both the concepts of dynamical causality and instantaneous coupling. Gao and Tian [5] proposed methods to learn Granger causality graphs based on conditional mutual information statistics and a permutation test. Structural learning of the Granger causality graphs are based on the conditional independence relations of random variables at different times. In order to test Granger causality between two time series, we need to compute the effect of every lag variables to current variables, which is computation-intensive and time-consuming.

Directed information theory deals with communication channels with feedback. Amblard and Michel [6] showed that measures built from directed information theory in networks can be used to assess Granger causality graphs of stochastic

* Corresponding author.
E-mail address: gaoweistat@163.com (W. Gao).

http://dx.doi.org/10.1016/j.physa.2017.05.035
0378-4371/

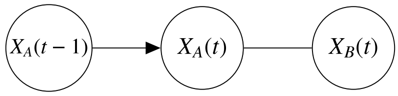

Fig. 1. A graph structure may cause spurious causality from X_A to X_B.

processes. Massey [7] showed that the appropriate information measure for directional dependence was no longer the mutual information, but the directed information. Kramer [8] accounted for the side information in the multivariate case and introduced causal conditioning. Quinn etc. [9–11] studied the relationships between Granger causality and the directed information theory and proposed the directed information graphs, but without considering the instantaneous dependence structure. Amblard [6] proposed an equivalent expression of Granger causality based on the directed information. Amblard and Michel [12,13] researched the links between the directed information theory [7,8] and the Granger causality graphs [4], provided a new insight into the problem of instantaneous coupling and showed how it affects the structure learning of graphical models. Amblard and Michel [12] derived a nonparametric estimator of directed information, which works efficiently for the graphical models when the dimension is not high.

As Amblard and Michel [12] presented, the instantaneous coupling strongly affects the edge detection in a graphical model. The general definition of Granger causality only takes into account the past variables, without considering the effect of the current variable, which will potentially cause spurious causality. Introducing the effect of the current variable need much increase of calculation amount and may difficult to obtain good interpretability. Then the existing research of directed information graph for Granger causality based on general definition.

Considering this problem, we propose a strict definition of Granger causality which includes the current variables in the conditional set. We prove that the directed information can be decomposed into three parts: a Granger causality term, an instantaneous coupling term and a term which caused by the autocorrelation. A measure based on directed information theory is applied as the statistics testing the existence of the Granger causality. The advantage is that directed information can measure the Granger Causality using only one statistics, while the other measures such as conditional mutual information requires p statistics, if the order of lag dependence is p. We learn the structure of the graphs from observed series based on the k-nn estimator for entropy and a permutation test. Finally, we run numerical simulations to demonstrate the efficiency of the proposed estimation methods.

The organization is as follows. A review of Granger causality and Granger causality graph is introduced in Section 2. Section 3 presents the decomposition theorem of directed information proposes the definition of Granger causality based on conditional directed information. Then in Section 4 we discuss the methods of structure learning of directed information graphs from observed series. In Section 5, we demonstrate the validity of the methods through simulation examples from linear and nonlinear models. Section 6 is the conclusion.

2. Granger causality

Suppose $X(t) = (X_1(t), X_2(t), \ldots, X_d(t))^T$, $t \in \mathbb{Z}$ is a $d-$ dimension stationary stochastic process defined on a probability space $(\Omega, \mathscr{F}, \mathbb{P})$. Let $V = \{1, 2, \ldots, d\}$ be a finite and nonempty set, $X_V = \{X_V(t), t \in \mathbb{Z}\}$ be a multivariate time series. For $a \in V$, X_a is the corresponding component in X_V. Likewise, for any $A \subset V$, X_A is the corresponding multivariate subprocess with component series $\{X_a, a \in A\}$. If subsets A, B, C are disjoint and $A \cup B \cup C = V$, we say that A, B, C form a partition of V. The information obtained by observing X_A up to time t is denoted as $X_A^t = \{\cdots, X_A(t-2), X_A(t-1), X_A(t)\}$.

The original definition of Granger causality says that conditioning on the past of the side information X_C and the past of X_B, $X_B(t)$ is not independent of the past of X_A.

Definition 1 (*[4]*). Suppose $X(t) = (X_1(t), X_2(t), \ldots, X_d(t))^T$, $t \in \mathbb{Z}$ is a $d-$ dimension stationary stochastic process, $V = \{1, 2, \ldots, d\}$, A, B, C form a partition of V. X_A does not Granger cause X_B relative to V if and only if $X_B(t) \perp\!\!\!\perp X_A^{t-1} | X_B^{t-1}, X_C^{t-1}, \forall t \in \mathbb{Z}$.

The definition above only contains the information of X_B^{t-1}, X_C^{t-1}, without including of the effect of the instantaneous variables $X_A(t)$ and $X_C(t)$ in the conditional set.

Example 2.1. Suppose series X_A, X_B have relations shown in Fig. 1. Definition 1 will result in spurious causality from $X_A(t-1)$ to $X_B(t)$ because $X_A(t-1) \not\perp\!\!\!\perp X_B(t)$, $X_A(t-1) \perp\!\!\!\perp X_B(t) | X_A(t)$.

In principle, causality must consider the directions of the time. Leaking the effect of the current variables may cause spurious causality. Considering this reason, we define a more strict version of Granger causality as follows.