Statistics and Probability Letters 142 (2018) 39–43

Contents lists available at ScienceDirect

Statistics and Probability Letters

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

Granger causality between vectors of time series: A puzzling property

Umberto Triacca

Department of Computer Engineering, Computer Science and Mathematics, University of L'Aquila, L'Aquila, Italy

ARTICLE INFO

Article history:
Received 22 November 2017
Received in revised form 18 June 2018
Accepted 20 June 2018
Available online 4 July 2018

Keywords:
Granger causality
Hilbert spaces
Time series

ABSTRACT

Let us consider a discrete-time n-dimensional stochastic process z, with components $x = (x_1, \ldots, x_{m_1})'$ and $y = (y_1, \ldots, y_{m_2})'$, $m_1 + m_2 = n$. We want to study causality relationships between the variables in x and$y$. Suppose that we find that y Granger causes x. Then we would expect to be able to pick out at least one of these variables, say y_j, having a causal impact on x. It turns out that, when we consider the conditioning information set defined by the past observations of x and all the y_i, $i \neq j$, it may be that y_j has no causal impact on x, irrespective of the particular $j = 1, 2, \ldots, m_2$ that we tried to pick out. This is a puzzling property. The paper provides a condition under which this property cannot hold.

1. Introduction

The concept of Granger causality was first introduced by Wiener (1956) and later reformulated and formalized by Granger (1969). A precise statement of this notion can be made in different ways. In this paper we will adopt the definition of noncausality used by Dufour and Renault (1998).

It is well known that the Granger causal links are sensitive to the information sets which are employed in the analysis. Changing the conditioning information sets, for example by extending or reducing the number of time series in the study, may lead to different Granger causal links. A causal statement is never intrinsic but always relative to given information sets. There is an extensive literature on these topics. Pertinent general theoretical discussions and developments are provided in: Lütkepohl (1982), Dufour and Renault (1998), Triacca (1998), Hosoya (2001), Giles (2002), Eichler (2007), Hill (2007) and Al-Sadoon (2014).

The contribution of this paper is to investigate a particular form of lack of invariance of causality statements to changes in the conditioning information sets. Let us consider a discrete-time n-dimensional stochastic process z, with components $x = (x_1, \ldots, x_{m_1})'$ and $y = (y_1, \ldots, y_{m_2})'$, $m_1 + m_2 = n$. We want to study causality relationships between the variables in x and y.

Suppose that we find that y Granger causes x. Then we would expect to be able to pick out at least one of these variables, say y_j, having a causal impact on x. It turns out that, when we consider the conditioning information set defined by the past observations of x and all the y_i, $i \neq j$, it may be that y_j has no causal impact on x, irrespective of the particular $j = 1, 2, \ldots, m_2$ that we tried to pick out. This kind of counter-intuitive result reveals some flaw in our way to tackle the causality analysis. Generally, the causal structure of a multivariate stochastic process does not allow conclusions concerning the causal structure of its subprocesses. Formally, this puzzling property has been established by Dufour and Renault (1998, p. 1105). Recently, Renault and Triacca (2015) have provided a condition (of global separability), regarding the causal variables $y = (y_1, \ldots, y_{m_2})'$, under which this property does not hold and hence the causal structure of a multivariate

E-mail address: umberto.triacca@univaq.it.

https://doi.org/10.1016/j.spl.2018.06.009

stochastic process depends on the causal structure of its subprocesses and vice versa. However, a problem with this notion of global separability is that it appears to be a high-level assumption without great intuitive underpinning.

The goal of this paper is to find a condition sufficient for the global separability, but with a more primitive interpretation. In particular, we will prove that the global separability is implied by a condition concerning the distance between the information sets used in the definition of Granger causality.

The set-up used in this paper borrows heavily from Dufour and Renault's (1998) framework for Granger causality. Thus, the presented results are obtained for processes in L^2 (i.e., processes with finite second moments), without any assumption on stationarity or specific parametric forms. However, it is important to underline that working in L^2 has enormous implications for causal structures. The findings are intimately linked to the structure of Hilbert spaces. That is, working in L^2 without restrictions on stationarity or function form is still a major restriction (see Hill (2006)).

The paper is organized as follows. In Section 2, we introduce the necessary notation and definitions. The main results are presented in Section 3.

2. Notation and definitions

Let $L^2 = L^2(\Omega, \mathcal{A}, P)$ be the Hilbert space of the real square-integrable random variables, defined on common probability space $(\Omega, \mathcal{A}, P)$ and let $\mathbb{E}$ be the expectation operator in this space. We define the inner product by $\langle u, v\rangle = \mathbb{E}(uv)$ for all $u, v \in L^2$ and the norm to be $\|u\|^2 = \langle u, u\rangle$ for all $u \in L^2$. A non-empty subset M of L^2 is called a (linear) manifold if for all $u, v \in M$ and all scalars α_1, α_2, we have $\alpha_1 u + \alpha_2 u \in M$. M is called a *subspace*, if it is a closed manifold, that is it contains the limit of every Cauchy sequence in M. The smallest subspace of L^2 which contains M, denoted by $\overline{\text{span}}(M)$, is called subspace spanned by M. For any two subspaces M and N of L^2, we will denote $M \vee N$ the subspace spanned by the elements of M and N, that is $M \vee N = \overline{\text{span}}(M \cup N)$. Further, following the mathematics literature, we use $M + N$ to denote the set of all elements of the form $u + v$ with $u \in M$ and $v \in N$. It is important to note that the vector sum of two subspaces can fail to be a subspace (Halmos, 1957, Section 8.6). The vector sum of a sequence $\{M_i\}$ of subspaces, in symbols $\sum_i M_i$, is the set of all vectors of the form $\sum_i u_i$ with $u_i \in M_i$ for all i. Let A and B be nonempty subsets of L^2. The distance between A and B, denoted $d(A, B)$, is defined as follows:

$$d(A, B) = \inf\{\|a - b\|\, ;\, a \in A, b \in B\}.$$

For a subspace M of L^2, its unit sphere, denoted $S(M)$, is the set

$$\{m \in M : \|m\| = 1\}.$$

Now, we consider a nondecreasing sequence I of subspaces $I(t)$ of L^2,

$$I = \{I(t);\, t \in \mathbb{Z}, t > \omega\} \quad \text{and} \quad t < t' \Rightarrow I(t) \subseteq I(t') \;\; \forall t > \omega,$$

where $\omega \in \mathbb{Z} \cup \{-\infty\}$ represents a 'starting point'. The 'starting point' ω is typically equal to a finite initial date (such as $\omega = -1, 0$ or 1) or to $-\infty$; $I(t)$ is said the 'reference information set'. Let x be an $m_1 \times 1$ vector process in L^2, that is

$$x = \{x(t);\, t \in \mathbb{Z}, t > \omega\},\;\; x(t) = \left(x_1(t), \ldots, x_{m_1}(t)\right)',\;\; x_i(t) \in L^2 \; (i = 1, \ldots, m_1).$$

We suppose that the information sequence I is conformable with x, that is

$$x(\omega, t] \subseteq I(t) \;\; \forall t > \omega$$

where $x(\omega, t]$ is the subspace spanned by the components $x_i(\tau)$, $i = 1, \ldots, m_1$ of $x(\tau)$, $\omega < \tau \leq t$. Let V be a subspace of L^2, for any positive integer h we denote $P\left[x_i(t+h)|V\right]$ the orthogonal projection of $x_i(t+h)$ on V and $P\left[x(t+h)|V\right] = \left(P\left[x_1(t+h)|V\right], \ldots, P\left[x_{m_1}(t+h)|V\right]\right)'$. Now, we consider an $m_2 \times 1$ vector process y in L^2, that is

$$y = \{y(t);\, t \in \mathbb{Z}, t > \omega\},\;\; y(t) = \left(y_1(t), \ldots, y_{m_2}(t)\right)',\;\; y_i(t) \in L^2 \; (i = 1, \ldots, m_2).$$

We will consider also the following subspaces of L^2: $y(\omega, t]$ the subspace spanned by the components $y_i(\tau)$, $i = 1, \ldots, m_2$ of $y(\tau)$, $\omega < \tau \leq t$; $y_j(\omega, t]$ the subspace spanned by $\left\{y_j(\tau);\, \omega < \tau \leq t\right\}$; $y^{(j)}(\omega, t]$ the subspace spanned by $\{y_i(\tau);\, \omega < \tau \leq t,\; i = 1, \ldots, m_2,\; i \neq j\}$.

Following Dufour and Renault (1998) we now give the following definitions of noncausality.

Definition 1 (*Noncausality at Different Horizons*)**.** For $h \in \mathbb{N}$, we say that: (i) y does not cause x at horizon h given I (denoted $y \overset{h}{\nrightarrow} x|I$) if

$$P(x(t+h)|I(t)) = P(x(t+h)|I(t) \vee y(\omega, t]) \;\forall t > \omega;$$

(ii) y_j does not cause x at horizon h given $I^{(j)}(t) = I(t) \vee y^{(j)}(\omega, t]$ (denoted $y_j \overset{h}{\nrightarrow} x|I^{(j)}$) if

$$P(x(t+h)|I^{(j)}(t)) = P(x(t+h)|I(t) \vee y(\omega, t]) \;\forall t > \omega;$$