Article

Non-Causality Due to Included Variables

Umberto Triacca (ID)

Department of Information Engineering, Computer Science and Mathematics University of L'Aquila,
67100 Coppito, Italy; umberto.triacca@univaq.it; Tel.: +39-064456100

Abstract: 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. Consider a discrete-time three-dimensional stochastic process $z = (x, y_1, y_2)'$. We want to study causality relationships between the variables in $y = (y_1, y_2)'$ and x. Suppose that in a bivariate framework, we find that y_1 Granger causes x and y_2 Granger causes x, but these relationships vanish when the analysis is conducted in a trivariate framework. Thus, the causal links, established in a bivariate setting, seem to be spurious. Is this conclusion always correct? In this note, we show that the causal links, in the bivariate framework, might well not be 'genuinely' spurious: they could be reflecting causality from the vector y to x. Paradoxically, in this case, it is the non-causality in trivariate system that is misleading.

Keywords: Granger causality; Hilbert spaces; time series

JEL Classification: C10; C32

1. Introduction

Following a suggestion of Wiener (1956), Granger (1969) introduced a concept of causality in a time series framework. Granger's definition concerns the predictability of a given stochastic process. Namely, if a process y contains information in the past terms that helps in the prediction of the process x, and if this information is contained in no other process used in the predictor, then y is said to cause x. This notion of Granger causality is used here. It is well-known that Granger non-causality, in a bivariate system, may be due to an omitted variable (see Lutkepohl (1982)). In this paper, we consider the opposite situation, which is apparent non-causality due to an included variable.

Consider a discrete-time three-dimensional stochastic process $z = (x, y_1, y_2)'$, and consider causality relationships between the variables in $y = (y_1, y_2)'$ and x. Suppose that in a bivariate framework, we find that y_1 Granger causes x and y_2 Granger causes x, but these relationships vanish when the analysis is conducted in a trivariate framework. Thus, the causal links, established in a bivariate setting, seem to be spurious. Is this conclusion always correct? In this note, we show that the causal links, in a bivariate framework, might not be 'genuinely' spurious: they could be reflecting causality from the vector y to x. Paradoxically, in this case, it is the non-causality in the trivariate system that is misleading. A finding of non-causality can thus lead to the potentially false inference of non-causation where causation is present—i.e., to a spurious conclusion of non-causality.

The rest of the paper is organized as follows. Section 2 introduces the definitions of causality used throughout the paper. Section 3 presents some theoretical results. A prescriptive implication is reported in Section 4. Finally, Section 5 presents the conclusions.

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 E be the expectation operator in this space. We define the inner product to be $\langle u, v \rangle = E(uv)$ 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 that 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, which is $M \vee N = \overline{\text{span}}(M \cup N)$. Now, we consider a non-decreasing sequence I of subspaces $I(t)$ of L^2,

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

where $\omega \in \mathbb{Z} \cup \{-\infty\}$. $I(t)$ is called the 'reference information set'. Let x be a scalar process in L^2. We suppose that the information sequence I is conformable with x, which is

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

where $x(\omega, t]$ is the subspace spanned by $x(\tau)$, $\omega < \tau \leq t$. Let V be a subspace of L^2, and, for any positive integer h, we denote $P[x(t+1)|V]$ the orthogonal projection of $x(t+1)$ on V. Now, we consider a 2×1 vector process y in L^2, which is

$$y = \{y(t); t \in \mathbb{Z}, t > \omega\}, \ y(t) = (y_1(t), y_2(t))', \ y_i(t) \in L^2 \ (i = 1, 2).$$

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

Now, we can give the following definitions of non-causality.

Definition 1. *y does not cause x given I (denoted $y \nrightarrow x | I$) if*

$$P(x(t+1)|I(t)) = P(x(t+1)|I(t) \vee y(\omega, t]) \ \forall t > \omega.$$

Definition 2. *(Bivariate non-causality) y_j does not cause x given I (denoted $y_j \nrightarrow x | I$) if*

$$P(x(t+1)|I(t)) = P(x(t+1)|I(t) \vee y_j(\omega, t]) \ \forall t > \omega.$$

Definition 3. *(Trivariate non-causality) y_j does not cause x given $I^{(j)}(t) = I(t) \vee y^{(j)}(\omega, t]$ (denoted $y_j \nrightarrow x | I^{(j)}$) if*

$$P(x(t+1)|I^{(j)}(t)) = P(x(t+1)|I(t) \vee y(\omega, t]) \ \forall t > \omega.$$

3. Results

In this section, we show that if there is a causal link from the vector y to x ($y \rightarrow x | I$) and there is non-causality in trivariate system ($y_1 \nrightarrow x | I^{(1)}$ and $y_2 \nrightarrow x | I^{(2)}$), then we necessarily find causality in the bivariate framework ($y_1 \rightarrow x | I$ and $y_2 \rightarrow x | I$). In order to do this, first, we present the following propositions.

Proposition 1. *If $y_i \nrightarrow x | I^{(i)}$ and $y_j \nrightarrow x | I$ with $i \neq j$, then $y \nrightarrow x | I$.*

Proof. Without loss of generality, we assume that $y_1 \nrightarrow x | I^{(1)}$ and $y_2 \nrightarrow x | I$. This implies

$$P(x(t+1) | I^{(1)}(t)) = P(x(t+1) | I^{(1)}(t) \vee y_1(\omega, t]) \; \forall t > \omega$$

and

$$P(x(t+1) | I(t)) = P(x(t+1) | I(t) \vee y_2(\omega, t]) \; \forall t > \omega.$$

Now, we note that

$$I^{(1)}(t) = I(t) \vee y_2(\omega, t])$$

and

$$I^{(1)}(t) \vee y_1(\omega, t] = I(t) \vee y(\omega, t]).$$

It follows that

$$P(x(t+1) | I(t)) = P(x(t+1) | I(t) \vee y(\omega, t]) \; \forall t > \omega,$$

that is, $y \nrightarrow x | I$. □

Proposition 2. *If* $y \rightarrow x | I$, $y_1 \nrightarrow x | I^{(1)}$ *and* $y_2 \nrightarrow x | I^{(2)}$, *then* $y_1 \rightarrow x | I$ *and* $y_2 \rightarrow x | I$.

Proof. If $y \rightarrow x | I$, by Proposition 1, we have $y_1 \rightarrow x | I^{(1)}$ or $y_2 \rightarrow x | I$ and $y_2 \rightarrow x | I^{(2)}$ or $y_1 \rightarrow x | I$. Since, by hypothesis, we have that $y_1 \nrightarrow x | I^{(1)}$ and $y_2 \nrightarrow x | I^{(2)}$, it follows that $y_1 \rightarrow x | I$ and $y_2 \rightarrow x | I$. □

Proposition 2 tells us that if there is a causal link from the vector y to x and if the components y_1 and y_2 are 'redundant' variables, i.e., y_1 provides the same useful information as y_2, for forecasting x, then necessarily $y_1 \rightarrow x | I$ and $y_2 \rightarrow x | I$. The causal links, in the bivariate system, must not be interpreted as spurious: in this case, they reflect the causal link from the vector y to x.

4. A Prescriptive Implication

It is well known that the omission of relevant variables may induce spurious apparent Granger causality among the remaining variables. In order to address this problem, the common practice suggests to enlarge the set of variables. However, our result makes it clear that this practice may be dangerously misleading under certain conditions. Thus, a wider information set is not necessarily always an unalloyed blessing in the context of binary causal relations: a problem of apparent non-causality due to redundant included variables could be present. For example, let us consider a situation where we find that y_1 Granger causes x and y_2 Granger causes x, in a bivariate setting. Furthermore, suppose that these causal links vanish when we consider the trivariate system (x, y_1, y_2). Against this backdrop, before concluding that the finding of causality is spurious, the results derived here indicate that one should investigate the possibility of a causal relationship between the vector (y_1, y_2) and x. Only if one finds that (y_1, y_2) does not cause x can one conclude that the causal links found in the bivariate setting are genuinely spurious.

5. Conclusions

Generally, the causal structure of a multivariate stochastic process does not allow conclusions concerning the causal structure of its sub process. Causal statements made in a trivariate setting will not necessarily retain their validity in a bivariate framework. A bivariate analysis can exclude relevant variables and lead to spurious correlation, and thereby to erroneous conclusions. However, Proposition 2 says to us that the causal links in the bivariate setting ($y_1 \rightarrow x | I$ and $y_2 \rightarrow x | I$) could not be spurious: they could reflect the causality from the vector $(y_1, y_2)'$ to x ($y \rightarrow x | I$). In this framework, it is the non-causality in the trivariate system ($y_1 \nrightarrow x | I^{(1)}$ and $y_2 \nrightarrow x | I^{(2)}$) that is spurious. This happens because none of the components of $(y_1, y_2)'$ contain unique information about x that is not available in the other. Therefore, each variable can act as a proxy for the other one, and eliminating one of them from the system does not reduce the quality of forecasts.

The determination of an appropriate information set leads to one common source of difficulty in a Granger non-causality study. When a bivariate system is used, the causal links could be spurious. On the other hand, when the analysis is conducted in a trivariate system, the findings of non-causality could be misleading because a problem of non-causality due to redundant included variables could be present.

Conflicts of Interest: The author declares no conflict of interest.

References

Granger, Clive William John. 1969. Investigating Causal Relations by Econometric Models and Cross-Spectral Methods. *Econometrica* 37: 424–38.

Lutkepohl, Helmut. 1982. Non-causality due to omitted variables. *Journal of Econometrica* 19: 367–78.

Wiener, Norbert. 1956. The Theory of Prediction. In *Modern Mathematics for Engineers*. Edited by Edwin F. Beckenback. New York: McGraw-Hill, chp. 8.