ARTICLE IN PRESS

Journal of Econometrics ▮ (▮▮▮▮) ▮▮▮–▮▮▮

Contents lists available at ScienceDirect

Journal of Econometrics

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

Geometric and long run aspects of Granger causality☆

Majid M. Al-Sadoon*

Universitat Pompeu Fabra & Barcelona GSE, Spain

ARTICLE INFO

Article history:
Received 8 October 2009
Received in revised form
6 August 2013
Accepted 19 August 2013
Available online xxxx

JEL classification:
C10
C32
C51

Keywords:
Granger causality
Long run Granger causality
L^2-mean-reversion
ρ-mixing
Cointegration
VARMA
Controllability
Kalman decomposition
Linear rational expectations

ABSTRACT

This paper extends multivariate Granger causality to take into account the subspaces along which Granger causality occurs as well as long run Granger causality. The properties of these new notions of Granger causality, along with the requisite restrictions, are derived and extensively studied for a wide variety of time series processes including linear invertible processes and VARMA. Using the proposed extensions, the paper demonstrates that: (i) mean reversion in L^2 is an instance of long run Granger non-causality, (ii) cointegration is a special case of long run Granger non-causality along a subspace, (iii) controllability is a special case of Granger causality, and finally (iv) linear rational expectations entail (possibly testable) Granger causality restriction along subspaces.

1. Introduction

First suggested by Wiener (1956) and later developed by Granger (1969), Granger causality (GC) and Granger non-causality (GNC) are two of the most important concepts of time series econometrics. Many extensions have been proposed throughout the years: multivariate time series (Tjøstheim, 1981), enlarged information sets (Hsiao, 1982), variable horizons (Dufour and Renault, 1998), etc.[1] Yet problems of interpretation have plagued it since its inception (see e.g. Hamilton, 1994) and some have argued that it may fail to capture what is actually meant by causality (see Hoover, 2001 or Pearl, 2009). Against this backdrop, the purpose of this paper is to demonstrate that GC is a much deeper concept than previously thought, going to the heart of many other concepts in time series analysis. This is done without taking any particular stance on the philosophical or empirical applicability of GC per se. Suffice it to say that GC remains an important element of causal analysis in a dynamic setting and that it does capture structural causality under certain conditions (White and Lu, 2010; White et al., 2010; White and Pettenuzzo, 2011; White et al., 2012). In such instances it makes sense to use causal language such as "cause" and "effect" in referring to variables associated by GC and we will on occasion do so in this paper with the understanding that those conditions are met.

This paper proposes two extensions to Dufour and Renault (1998) (DR): (i) it takes into account the subspaces of GC and (ii) it considers long run GC. To motivate the first extension, suppose that X and Y are multivariate processes and Y Granger-causes X. Now it may be that variations in X along some directions cannot be attributed to Y. Likewise, it may be that certain linear combinations of Y do not help predict X. Thus standard multivariate GNC tests may not give the full picture of the dependence structure. To motivate the second extension, frequency-domain results are available for checking long run GNC (Hosoya, 1991, 2001).

☆ An earlier draft of this paper was titled, "Causality along subspaces: Theory." My most sincere thanks and gratitude go to Sean Holly for his help, support, and patience throughout the writing of this paper, which was part of the author's Ph.D. Thesis at the University of Cambridge. Thanks also go to M. Hashem Pesaran, Hal White, Richard J. Smith, George Kapetanios, Rod McCrorie, Cheng Hsiao, two anonymous referees, and to seminar participants at the University of Cambridge and the University of St. Andrews. This paper is dedicated to the memory of Halbert L. White, Jr. He is profoundly missed. Financial support of the Yousef Jameel scholarship is gratefully acknowledged.

* Tel.: +34 93 542 2682.
E-mail address: majid.alsadoon@gmail.com.

[1] Excellent surveys can be found in Geweke (1984), Hamilton (1994), and Lütkepohl (2006).

0304-4076/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.jeconom.2013.08.019

There are also time-domain results for cointegrated VAR models (Granger and Lin, 1995; Bruneau and Jondeau, 1999; Yamamoto and Kurozumi, 2006). It would be useful to obtain time-domain criteria for long run GNC for a wider class of processes.

Based on the aforementioned extensions, it is shown that: (i) L^2-mean-reversion, a weaker form of weak dependence than ρ-mixing, is an instance of long run GNC, (ii) cointegration is a special case of long run GNC along a subspace, (iii) controllability is a special case of subspace GC and Kalman's controllability decomposition is a partial converse of a result by DR, and finally (iv) linear rational expectations entail (possibly testable) GNC restriction along subspaces. Additionally, the paper presents extensions of various results by DR to subspace GNC in linear L^2 processes, including VARMA.

Now GC has been known to be associated with cointegration, controllability, and rational expectations equilibria for quite some time now. However these links have been established in rather restrictive contexts and do not span the full extents of the relationships. In particular, the association with cointegration was known to hold only in the context of bivariate models (Granger, 1988b), whereas we shall see that cointegration is a particular form of long run subspace GNC in any multivariate L^2 process. The association with controllability, on the other hand, was only shown in rather extreme forms of optimal control, where the policymaker cares only about a single variable in the model (Granger, 1988a). We will see that controllability in its most general form (see e.g. Kailath, 1980) is a particular instance of subspace GC. Finally, the association with rational expectations has been explored by Hansen and Sargent (1980) although in the highly specialized context of stochastic linear–quadratic control. They find that GC determines which variables ought to enter into the decision rule. In contrast, the result of this paper, which applies to a larger class of linear rational expectations model and any variable therein (whether or not it is a decision rule), is that the forward component of a rational expectations equilibrium lies within a particular subspace of GNC. It is important to emphasize that none of these results would have been possible without the two extensions proposed in this paper. The general theme of this paper, kindly noted by an anonymous referee, is therefore: "[Granger] causality is not invariant to linear projections onto alternative subspaces".

In addition to the above literature, various papers have considered time series dependence along subspaces. Velu et al. (1986) consider the problem of finding the subspaces along which a stationary VAR is forecastable. Otter (1990) considers the problem of finding the subspace of future variables predictable by past variables. Related to this work is Brillinger (2001), who considers the problem of approximating a time series X by a filter of Y where the filter is of reduced rank and both series are stationary. There are also a number of papers that have recently built on DR. Eichler (2007) uses DR's results to conduct a graph-theoretic analysis in light of recent advances in the artificial intelligence literature on causality (Pearl, 2009). Hill (2007) develops DR's results into a procedure for finding the exact horizon at which fluctuations in one variable anticipate changes in another variable when the model is trivariate. Dufour and Taamouti (2010) develop measures of GC at finite horizons.

The paper proceeds as follows. Section 2 reviews some Hilbert space theory and sets the notation. Section 3 develops the main concepts of subspace and long run GNC as extensions to DR. Section 4 considers long run GNC in more detail. Section 5 specializes the theory to linear invertible processes. Section 6 specializes further to invertible VARMA processes. Section 7 considers the connection to controllability. Section 8 considers the connection to linear rational expectations equilibria. Section 9 concludes and the last section is an Appendix.

2. Review of Hilbert space theory and notation[2]

Throughout this paper, we work with a single probability space $(\Sigma, \mathscr{F}, \mathbb{P})$ with $\mathbb{E}$ as the expectation operator. We define L^2 to be the Hilbert space of random variables with finite second moments. The inner product is defined as $\langle X, Y\rangle = \mathbb{E}(XY)$ for all $X, Y \in L^2$ (we reserve $\|\cdot\|$ for the Euclidean vector norm and the norm it induces on matrices). We abuse notation by considering a random vector to be in L^2 if all its elements are in L^2. For a σ-algebra $\mathscr{X} \subseteq \mathscr{F}$, we take $L^2(\mathscr{X})$ to be the space of $\mathscr{X}$-measurable random variables in L^2.

If H and G are subspaces of L^2 then define $H + G = \overline{\text{sp}}\{H, G\}$, the closure of the span of all linear combinations of the elements of G and H.[3] We set $H - G = H \cap G^{\perp}$, the part of H orthogonal to G. This subspace is closed whenever H is closed and is defined even when $G \cap H = \{0\}$, in which case $H - G = H$.

The time indexing set will be $(\omega, \infty) \subseteq \mathbb{Z}$ with $\omega \in \{-\infty\} \cup \mathbb{Z}$ for all processes in this paper. The information or history at time $t > \omega$ is denoted by $I(t)$, a closed subspace of L^2 satisfying $I(t) \subseteq I(t')$ whenever $\omega < t \leq t'$. $I = \{I(t) : t > \omega\}$ is an information set. If X is an n-dimensional stochastic process in L^2 then for $\omega \leq t < t'$ define, $X(t, t'] = \overline{\text{sp}}\{X_{is} : t < s \leq t', 1 \leq i \leq n\}$. $X[t, t')$ is defined in similar fashion for $\omega < t \leq t'$ and so is $X[t, t']$. The information set I is said to be conformable with X if $X(\omega, t] \subseteq I(t)$ for all $t > \omega$. The most frequently encountered information sets in this paper take the form, $I(t) = H + X(\omega, t]$ for all $t > \omega$ for some L^2 random vector process X, where $H \subseteq L^2$ is a closed subspace. When $I(t) = H$ for all $t > \omega$ we will refer to I as H. The remote information set of X is defined as $\bigcap_{t>\omega} X(\omega, t]$. We will also require $\mathscr{X}_T \subseteq \mathscr{F}$, the σ-algebra generated by $\{X(t) : t \in T\}$, where T is a subset of the time indexing set (ω, ∞).

If H is a closed subspace of L^2 then the orthogonal projection of $X \in L^2$ onto H (or the best linear predictor of X given H) is denoted by $P(X|H)$. If X is a vector of n variables in L^2 then $P(X|H) = (P(X_1|H), \ldots, P(X_n|H))'$.

3. Cartesian and subspace Granger causality

First, we consider the basic idea behind subspace and long run GC as an extension to Cartesian GC. This study is conducted within a large class of time series processes, namely L^2 processes.

We take the following to be our most basic assumption.

Assumption 1. Let $\omega, \varpi \in \{-\infty\} \cup \mathbb{Z}$ and $\omega \leq \varpi$. $X = \{X(t) : \omega < t < \infty\}$ and $Y = \{Y(t) : \omega < t < \infty\}$ are L^2 processes, of finite dimensions n_X and n_Y respectively. We also take I to be an information set. Here ω specifies the start of the process and ϖ specifies the start of the prediction period. □

We will be interested in studying the predictability of X in terms of Y in the context of information set I. Because prediction sometimes requires initial conditions to be specified, this predictability is assessed over a range of periods (ϖ, ∞), which may be a proper subset of the time indexing set, (ω, ∞).[4] Typically, I is assumed to include all the variables that may help predict X, including X and

[2] Excellent overviews of the applications of Hilbert space theory to time series analysis can be found in Brockwell and Davis (1991) and Pourahmadi (2001). This paper closely follows the notation of DR.

[3] The statistical literature uses "+" to refer to the linear span. However, DR use "+" to signify the closed linear span and we follow their notation. The two are not equivalent as demonstrated in Example 9.6 of Pourahmadi (2001).

[4] DR and Al-Sadoon (2009b) did not make this distinction clear. They derive their general L^2 results for the case $\omega = \varpi$ but when discussing linear invertible models they allow for $\omega < \varpi$. This then begs the question of whether their general L^2 results continue to hold for linear invertible processes with initial conditions.