Journal of Neuroscience Methods 201 (2011) 404–419

Contents lists available at ScienceDirect

Journal of Neuroscience Methods

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

Behaviour of Granger causality under filtering: Theoretical invariance and practical application

Lionel Barnett*, Anil K. Seth

Sackler Centre for Consciousness Science and School of Informatics, University of Sussex, Brighton BN1 9QJ, UK

ARTICLE INFO

Article history:
Received 2 June 2011
Received in revised form 1 August 2011
Accepted 4 August 2011

PACS:
87.19.L–
87.19.lj
87.10.Mn

Keywords:
Granger causality
Digital filtering
Vector autoregressive modelling
Time series analysis

ABSTRACT

Granger causality (G-causality) is increasingly employed as a method for identifying directed functional connectivity in neural time series data. However, little attention has been paid to the influence of common preprocessing methods such as filtering on G-causality inference. Filtering is often used to remove artifacts from data and/or to isolate frequency bands of interest. Here, we show [following Geweke (1982)] that G-causality for a stationary vector autoregressive (VAR) process is fully invariant under the application of an arbitrary invertible filter; therefore filtering cannot and does not isolate frequency-specific G-causal inferences. We describe and illustrate a simple alternative: integration of frequency domain (spectral) G-causality over the appropriate frequencies ("band limited G-causality"). We then show, using an analytically solvable minimal model, that in practice G-causality inferences often do change after filtering, as a consequence of large increases in empirical model order induced by filtering. Finally, we demonstrate a valid application of filtering in removing a nonstationary ("line noise") component from data. In summary, when applied carefully, filtering can be a useful preprocessing step for removing artifacts and for furnishing or improving stationarity; however filtering is inappropriate for isolating causal influences within specific frequency bands.

1. Introduction

A key theme in contemporary neuroscience is to move from localisation of function to characterisation of functional networks. In particular, analysis methods aimed at extracting directed functional (i.e., causal) connectivity from neural signals are increasingly in demand.[1] G-causality analysis is widely employed to identify causal connectivity in neural time series data. G-causality is a statistical measure of causality based on precedence and predictability. Put simply, if a variable A contains information that helps predict another variable B, better than can be done knowing only the past of B itself, then A is said to "G-cause" B. The concept has typically been operationalised in the context of linear VAR models and its uptake within neuroscience has been facilitated by the appearance of dedicated software toolboxes implementing the methods (Seth, 2010; Cui et al., 2008). However, the interaction of G-causality with standard data preprocessing procedures is not well understood and presents a possibly serious confound to many applications. In this paper, we focus on the effects of (temporal) filtering on G-causality. This is a crucial issue since filtering is often applied semi-automatically as a preprocessing step in many analyses. Most applications of filtering attempt to achieve one (or both) of two objectives: (i) removal of artifacts such as electrical line noise and (non-neural) physiological influences, and (ii) isolation of effects within a specific frequency band [e.g., the beta or gamma ranges in M/EEG (Pollonini et al., 2010; Wilson and Yan, 2010)]. Anticipating our results, we show that G-causality is theoretically invariant under the application of arbitrary (invertible) multivariate filters, and so cannot achieve the second objective. However, the invariance holds strictly for *stationary* data—stationarity being a prerequisite for G-causality analysis—so that filtering can be useful for artifact removal if it is able to render a previously nonstationary time series stationary. In practice, filtering can pose challenges for the effective estimation of the autoregressive models on which G-causality is based, hence the need for its careful application in the context of achieving or improving stationarity. Although our analysis is targeted at "explicit" filtering imposed by an experimenter as

Abbreviations: G-causality, Granger causality; iid, identically and independently distributed; MVGC, multivariate Granger causality; VAR, vector autoregressive; VMA, vector moving average; VARMA, vector autoregressive moving average; FIR, finite impulse response; IIR, infinite impulse response; OLS, ordinary least squares; AIC, Akaike information criterion; BIC, Bayesian information criterion; CV, cross-validation; EEG, electroencephalography; MEG, magnetoencephalography; fMRI, functional magnetic resonance imaging; BOLD, blood oxygen level dependent; HRF, hemodynamic response function; DTF, directed transfer function; PDC, partially directed coherence.

* Corresponding author. Tel.: +44 1273 699246.
E-mail addresses: l.c.barnett@sussex.ac.uk (L. Barnett), a.k.seth@sussex.ac.uk (A.K. Seth).

[1] We prefer the term *causal connectivity*, a description of the data, to *effective connectivity*, which implies a model of the underlying mechanism; see Bressler and Seth (2011).

0165-0270/$ – see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.jneumeth.2011.08.010

a data (pre)processing stage, our results may also have implications for "implicit" filtering that may arise as a result of physiological processes intervening between neural variables and observables, for example as manifest in the hemodynamic BOLD signal measured using fMRI.

In his seminal 1982 paper Geweke (1982) noted, but did not justify or explore, the invariance of G-causality under filtering via the somewhat oblique aside "[G-causality] is invariant with respect to scaling of X and Y; in fact it remains unchanged if X and Y are pre-multiplied by different invertible lag operators." Perhaps because there is no explicit reference to "filtering" this note appears to have been overlooked as G-causality has been taken up within neuroscience. More recently, researchers have worried that filtering does in fact affect G-causality (Florin et al., 2010; Seth, 2010). A recent study by Florin et al. (2010) suggested that application of filtering to neural data disturbs the information content and time ordering of the data, leading to spurious and missed causalities (Type I and Type II errors, respectively). Their conclusion is based on the correct observation that filtering in general alters the regression coefficients of VAR models of the data. They then show using numerical simulations that filtering induces Type I and Type II errors in sample.[2] However, they did not make any analytical connection between the two observations. In fact, as we argue, the errors observed in simulation by Florin et al. derive from the difficulties inherent in fitting VAR models to filtered data, not from the filtering process *per se*. In particular, filtering generally induces a large increase in the empirical model order (the number of lagged observations incorporated into a VAR), leading to model mis-specification given limited data.

Our paper is organised as follows: in Section 2 we define G-causality in both the time and frequency domains, for unconditional and conditional situations, and for both univariate and multivariate (block, ensemble) variables. We also discuss estimation for finite-sample empirical data and significance testing. Readers familiar with the mathematical basis of G-causality may wish to skip this section, referring to it where needed for notation. In Section 3 we demonstrate analytically the invariance of G-causality under the application of an (almost) arbitrary stable, invertible, multivariate filter. The invariance is completely general, applying to all the varieties of G-causality just mentioned. We then consider issues arising in empirical estimation of G-causality, suggesting several reasons why filtering may corrupt empirical estimates despite the theoretical invariance. As mentioned, these turn principally on an increase in empirical model order induced by filtering; filtering may also induce near-nonstationarity and other numerical instabilities. Consequently, we argue that (i) filtering can be useful for preprocessing nonstationary (or near-nonstationary) time series and (ii) estimation of G-causality within specific frequency bands can be accomplished by integrating the frequency domain G-causality over the appropriate frequencies ("band limited G-causality"). Section 4 introduces a minimal VAR system for which G-causalities can be obtained analytically. We use this model to test how empirical estimates of G-causality are influenced by both FIR and IIR filters. We compare estimates of model order for unfiltered and filtered processes, showing a large increase in optimal (empirical) model order following filtering, as well as an increase in the likelihood of unstable VAR models. We then analyse the effects of model order and filtering on statistical significance testing, showing [consistent with Florin et al. (2010); Seth (2010)] increases in both Type I and Type II errors after filtering. We explain this result by showing a strong association between increased error rates and an increase in VAR model order entailed by filtering. Based on these findings, we demonstrate a useful example of filtering to remove line noise. Finally, we show that band-limited G-causality on unfiltered data correctly identifies frequency specific causal interactions, whereas G-causality on filtered data does not. Our conclusions are summarised and discussed in Section 5.

[2] A similar corruption of G-causality inferences by filtering was shown in another set of recent simulations (Seth, 2010).

2. Multivariate G-causality (MVGC)

Consider a covariance-stationary, n variable, VAR(p) process $\boldsymbol{U}_t$ (the "universe" of measurable variables) specified by the model[3]

$$\sum_{k=0}^{p} A_k \cdot \boldsymbol{U}_{t-k} = \boldsymbol{\varepsilon}_t \tag{1}$$

for $-\infty < t < \infty$, where the $n \times n$ square matrices A_k, $k = 0, 1, 2, \ldots, p$ are the regression coefficients with $A_0 \equiv I$, the identity matrix, and $\boldsymbol{\varepsilon}_t$ are serially uncorrelated iid residuals (white noise) with covariance matrix $\Sigma \equiv \mathrm{cov}(\boldsymbol{\varepsilon}_t)$. We allow the model order p to be infinite. Introducing the lag operator $\mathcal{L}$ so that $\mathcal{L}\boldsymbol{U}_t = \boldsymbol{U}_{t-1}$, $\mathcal{L}^2\boldsymbol{U}_t = \boldsymbol{U}_{t-2}$, etc., we can write (1) in the form

$$A(\mathcal{L}) \cdot \boldsymbol{U}_t = \boldsymbol{\varepsilon}_t \tag{2}$$

where the p th order square matrix polynomial $A(z)$ is defined to be $A(z) \equiv \sum_{k=0}^{p} A_k z^k$, with $A(0) = I$.

Covariance-stationarity requires that $A(z)$ exists and is *invertible* for all z on the unit disk $|z| \leq 1$ in the complex z-plane (Hamilton, 1994); a VAR model of the form (2) is described as *stable* if it satisfies this condition. For the finite order case, this requires that all roots of the characteristic polynomial $\det(A(z^{-1}))$ lie strictly inside the unit circle. The maximum modulus of the roots of the characteristic polynomial is the *spectral radius* of the VAR model, written $\rho(A)$. Intuitively, $\rho(A)$ determines how rapidly autocorrelation of the VAR decays with increasing lag time, and stability requires that $\rho(A) < 1$.

Since the VAR (2) is assumed covariance-stationary, by the Wold decomposition theorem (Hamilton, 1994) it may be written equivalently in VMA form as

$$\boldsymbol{U}_t = H(\mathcal{L}) \cdot \boldsymbol{\varepsilon}_t \tag{3}$$

where the *transfer function* $H(z)$ for the model is the rational matrix function defined by $H(z) \equiv A(z)^{-1}$. In general, the VMA representation will be of *infinite* order.

2.1. Time domain

We consider firstly *unconditional* G-causality. Suppose that $\boldsymbol{U}_t$ is decomposed into two jointly distributed, multivariate processes $\boldsymbol{U}_t = \begin{pmatrix} \boldsymbol{X}_t \\ \boldsymbol{Y}_t \end{pmatrix}$ with $\dim(\boldsymbol{X}) = k$ and $\dim(\boldsymbol{Y}) = l$, $k + l = n$. We wish to ascertain the causal effect of the variable $\boldsymbol{Y}$ on the variable $\boldsymbol{X}$; i.e., the G-causality $\mathcal{F}_{\boldsymbol{Y} \to \boldsymbol{X}}$.

We may decompose the autoregression (2) as

$$\begin{pmatrix} A_{xx}(\mathcal{L}) & A_{xy}(\mathcal{L}) \\ A_{yx}(\mathcal{L}) & A_{yy}(\mathcal{L}) \end{pmatrix} \cdot \begin{pmatrix} \boldsymbol{X}_t \\ \boldsymbol{Y}_t \end{pmatrix} = \begin{pmatrix} \boldsymbol{\varepsilon}_{x,t} \\ \boldsymbol{\varepsilon}_{y,t} \end{pmatrix} \tag{4}$$

with VMA representation

$$\begin{pmatrix} \boldsymbol{X}_t \\ \boldsymbol{Y}_t \end{pmatrix} = \begin{pmatrix} H_{xx}(\mathcal{L}) & H_{xy}(\mathcal{L}) \\ H_{yx}(\mathcal{L}) & H_{yy}(\mathcal{L}) \end{pmatrix} \cdot \begin{pmatrix} \boldsymbol{\varepsilon}_{x,t} \\ \boldsymbol{\varepsilon}_{y,t} \end{pmatrix} \tag{5}$$

[3] In all that follows, bold type indicates a vector quantity and upper-case type denotes either a matrix or a random variable, depending on context. Vectors are considered to be *column* vectors. The symbol $^{\mathrm{T}}$ indicates matrix transpose; an asterisk denotes the conjugate transpose of a (complex) matrix, and $\det(\cdot)$ denotes the determinant of a (square) matrix.