Granger Causality via Vector Auto-Regression Tuned for FMRI Data Analysis

G. Chen[1], J. P. Hamilton[2], M. E. Thomason[2], I. H. Gotlib[2], Z. S. Saad[1], and R. W. Cox[1]

[1]Scientific and Statistical Computing Core, NIMH, National Institutes of Health, Bethesda, MD, United States, [2]Mood and Anxiety Disorders Laboratory, Department of Psychology, Stanford University, Stanford, CA, United States

Introduction: In addition to localizing task-related cortical and subcortical activity, FMRI data can be used to model various forms of interactions among active regions of interest (ROIs) in the brain. Granger causality (GC) is a connectivity modeling approach with the following premises: (a) a region causally connected to another would leave in the latter a time-lagged version of its own signal components; and (b) the latency in the signal indicates a causal relation. In a typical multivariate GC analysis, the causality among the pre-selected ROIs is determined in a vector auto-regressive (VAR) model with appropriate lags (time delay order) over which connectivity is to be examined. A second-stage analysis can then be conducted to examine group differences in connectivity. We present a multivariate implementation of GC that takes into account the idiosyncrasies of FMRI data and provides for inference on the sign of correlation, evolution of the network over multiple lags, model order tuning, and model diagnosis. We also propose a valid group analysis per lag based on *signed* path coefficients to reveal the network at the group level.

Methods: We start with an extended VAR(p) model $Y(t)=\alpha+A_1Y(t\text{-}1)+\ldots+A_pY(t\text{-}p)+BZ(t)+\varepsilon(t)$, where: $Y(t)$ is an $n\times1$ vector containing the BOLD response of n ROIs at time t; $Z(t)$ is an $q\times1$ vector containing exogenous variables (covariates or confounds) at time t; p and q are the number of lags and confounds respectively; α is an $n\times1$ vector modeling the baseline; A_i is an $n\times n$ matrix with each element reflecting the effect from one ROI to another at time lag i (i=1,…,p); B is an $n\times q$ matrix with each element modeling the effect of a confound on an ROI; and $\varepsilon(t)$ is the error term. The confounds can be baseline drift, head motion parameters, tasks of no interest, physiological measurements, or dummy variables corresponding to time breaks in the time series due to multiple blocks/runs/sessions and to time points censored because of signal irregularities.

Unlike previous implementations [1, 2] of the standard VAR(p) model $Y(t)=\alpha+A_1Y(t\text{-}1)+\ldots+A_pY(t\text{-}p)+\varepsilon(t)$ with significance based on F-statistics across multiple lags, our approach with an extended model allows us to perform whole network causality analysis while simultaneously modeling confounding effects, instead of regressing them out prior to the GC analysis. Furthermore, our implementation carries correlation signs at different lags all the way to the group level analysis, because reducing correlations to F-statistics and collapsing networks across lags can significantly alter the identified network. Four criteria (AIC, FPE, HQ, and SC) are provided to assist the user in selecting appropriate lag levels (order) of the model, and one network per lag is identified based on path coefficients and their t-statistics, potentially showing the network evolution through time (lags). Lastly, a set of model diagnosis tests are available to check for stationarity (testing roots of characteristic polynomial), residual normality tests (Gaussian process, skewness, and kurtosis), residual autocorrelation (portmanteau, Breusch-Godfrey, and Edgerton-Shukur tests), autoregressive conditional heteroskedasticity (ARCH) test for time-varying volatility, and structure stability (or stationarity detection). Our modeling approach was validated at the individual subject level by comparing with [2] when using the same model and input.

To illustrate the importance of preserving the correlation sign, we modeled the network among three ROIs—dorsolateral prefrontal cortex (DLPFC), dorsolateral caudate, and thalamus—that comprise the cognitive division of the cortico-striatal-thalamic loop, a well-validated neural circuit [3]. We collected 250 volume images during eye-closed rest [*cf.* ref. 4: 18 axial slices, spiral in-out; TR=1.2s, TE=30ms, FOV=22cm, matrix=64×64, slice thickness=5mm, flip angle=77°], from each of 20 participants.

Results: For this analysis, we adopted a lag of 1, as suggested by the model tuning step of our software written in R [5, 6]. The left half of Figure 1 shows the group network obtained with a one-sample t-test using log-transformed F-statistics of individual subjects for each connection, as typically practiced in the literature [*e.g.*, 1]. The network in the right half of Figure 1 was obtained with a one-sample t-test using the signed path coefficients instead. Note the marked differences between the two networks: none of the inter-regional connections were identified with transformed F-statistics as input, and the information about the correlation sign was totally lost, which can be attributed to either ignoring the path coefficient sign, the arbitrary transformation of the F-statistics, or both.

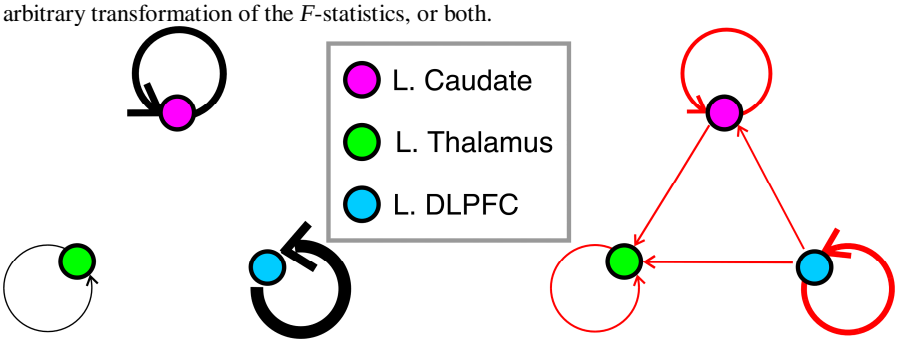

Figure 1 Two group networks of 3 ROIs in the left brain (DLPFC, dorsolateral caudate, and thalamus) were identified in a VAR(1) model. Confounds were 6 head motion parameters, respiration and heart rates, and 3rd-order Legendre polynomials to allow for slow drifts in the signal. Left and right networks were obtained with log-transformed F-statistics and signed correlation (respectively) from causality analysis with 20 subjects (p=0.05, corrected for multiple comparisons). Paths in red indicate positive correlation, while a black path means unknown correlation sign. Line thickness reflects the relative strength of a path. A circle around an ROI indicates a significant first-order auto-correlation.

Conclusions: We developed a platform-independent modeling tool that provides valid and consistent multivariate Granger causality analysis particularly suited for FMRI data. The program identifies patterns of association among brain ROIs that have been identified with other techniques, and generates a graphic representation of the identified network. Importantly, and unlike previous implementations of GC analysis for FMRI, this technique preserves information concerning the sign, in addition to the direction of prior prediction (causality) between ROI time courses from individual subjects all the way to the group level analysis. The software is in open-source R and available for download.

References

[1] A Roebroeck, E Formisano, and R Goebel, Mapping directed influence over the brain using granger causality mapping. NeuroImage, 25:230–242, 2005.
[2] AK Seth, Causal connectivity analysis of evolved neural networks during behavior. Network: Computation in Neural Systems. 16(1):35-54, 2005.
[3] GE Alexander, MR DeLong, PL Strick, Parallel Organization of Functionally Segregated Circuits Linking Basal Ganglia and Cortex. Ann Rev Neurosci, 9:357-381.
[4] GH Glover and CS Law, Spiral-in/out BOLD fMRI for increased SNR and reduced susceptibility artifacts. Magn Reson Med 46(3): 515–522, 2001.
[5] R Development Core Team, R: A language and environment for statistical computing. http://www.R-project.org. 2008.
[6] B Pfaff, Analysis of Integrated and Cointegrated Time Series with R. 2nd Edition. Springer, New York, 2008.