

Granger causality

 +12 Recommend this on Google+

Anil Seth (2007), Scholarpedia, 2(7):1667.

doi:10.4249/scholarpedia.1667

revision #91329 [link to/cite this article]

- **Anil Seth**, University of Sussex, UK

Granger causality is a statistical concept of causality that is based on prediction. According to Granger causality, if a signal X_1 "Granger-causes" (or "G-causes") a signal X_2 , then past values of X_1 should contain information that helps predict X_2 above and beyond the information contained in past values of X_2 alone. Its mathematical formulation is based on linear regression modeling of stochastic processes (Granger 1969). More complex extensions to nonlinear cases exist, however these extensions are often more difficult to apply in practice.

Granger causality (or "G-causality") was developed in 1960s and has been widely used in economics since the 1960s. However it is only within the last few years that applications in neuroscience have become popular.

Contents

- 1 Personal account by Clive Granger
- 2 Mathematical formulation
- 3 Spectral G-causality
- 4 Limitations and extensions
 - 4.1 Linearity
 - 4.2 Stationarity
 - 4.3 Dependence on observed variables
- 5 Application in neuroscience
- 6 Alternative techniques
 - 6.1 Information theory
 - 6.2 Maximum likelihood models
 - 6.3 Alternative frameworks
- 7 Concluding remarks
- 8 Resources
- 9 See also
- 10 External links
- 11 References
- 12 See Also

Personal account by Clive Granger

The following is a personal account of the development of Granger causality kindly provided by Professor Clive Granger (Figure 1). Please do not edit this section.

The topic of how to define causality has kept philosophers busy for over two thousand years and has yet to be resolved. It is a deep convoluted question with many possible answers which do not satisfy everyone, and yet it remains of some importance. Investigators would like to think that they have found a "cause", which is a deep fundamental relationship and possibly potentially useful.

In the early 1960's I was considering a pair of related stochastic processes which were clearly inter-related and I wanted to know if this relationship could be broken down into a pair of one way relationships. It was suggested to me to look at a definition of causality proposed by a very famous mathematician, Norbert Wiener, so I adapted this definition (Wiener 1956) into a practical form and discussed it.

Applied economists found the definition understandable and useable and applications of it started to appear. However, several writers stated that "of course, this is not real causality, it is only Granger causality." Thus, from the beginning, applications used this term to distinguish it from other possible definitions.

The basic "Granger Causality" definition is quite simple. Suppose that we have three terms, X_t , Y_t , and W_t , and that we first attempt to forecast X_{t+1} using past terms of X_t and W_t . We then try to forecast X_{t+1} using past terms of X_t , Y_t , and W_t . If the second forecast is found to be more successful, according to standard cost functions, then the past of Y appears to contain information helping in forecasting X_{t+1} that is not in past X_t or W_t . In particular, W_t could be a vector of possible explanatory variables. Thus, Y_t would "Granger cause" X_{t+1} if (a) Y_t occurs before X_{t+1} ; and (b) it contains information useful in forecasting X_{t+1} that is not found in a group of other appropriate variables.

Naturally, the larger W_t is, and the more carefully its contents are selected, the more stringent a criterion Y_t is passing. Eventually, Y_t might seem to contain unique information about X_{t+1} that is not found in other variables which is why the "causality" label is perhaps appropriate.

The definition leans heavily on the idea that the cause occurs before the effect, which is the basis of most, but not all, causality definitions. Some implications are that it is possible for Y_t to cause X_{t+1} and for X_t to cause Y_{t+1} , a feedback stochastic system. However, it is not possible for a determinate process, such as an exponential trend, to be a cause or to be caused by another variable.

It is possible to formulate statistical tests for which I now designate as G-causality, and many are available and are described in some econometric textbooks (see also the following section and the #references). The definition has been widely cited and applied because it is pragmatic, easy to understand, and to apply. It is generally agreed that it does not capture all aspects of causality, but enough to be worth considering in an empirical test.

There are now a number of alternative definitions in economics, but they are little used as they are less easy to implement.

Further references for this personal account are (Granger 1980; Granger 2001).



Figure 1: Prof. Clive W.J. Granger, recipient of the 2003 Nobel Prize in Economics

Mathematical formulation

G-causality is normally tested in the context of linear regression models. For illustration, consider a bivariate linear autoregressive model of two variables X_1 and X_2 :

$$\begin{aligned} X_1(t) &= \sum_{j=1}^p A_{11,j} X_1(t-j) + \sum_{j=1}^p A_{12,j} X_2(t-j) + E_1(t) \\ X_2(t) &= \sum_{j=1}^p A_{21,j} X_1(t-j) + \sum_{j=1}^p A_{22,j} X_2(t-j) + E_2(t) \end{aligned} \quad (1)$$

where p is the maximum number of lagged observations included in the model (the model order), the matrix A contains the coefficients of the model (i.e., the contributions of each lagged observation to the predicted values of $X_1(t)$ and $X_2(t)$), and E_1 and E_2 are residuals (prediction errors) for each time series. If the variance of E_1 (or E_2) is reduced by the inclusion of the X_2 (or X_1) terms in the first (or second) equation, then it is said that X_2 (or X_1) Granger-(G)-causes X_1 (or X_2). In other words, X_2 G-causes X_1 if the coefficients in A_{12} are jointly significantly different from zero. This can be tested by performing an F-test of the null hypothesis that $A_{12} = 0$, given assumptions of covariance stationarity on X_1 and X_2 . The magnitude of a G-causality interaction can be estimated by the logarithm of the corresponding F-statistic (Geweke 1982). Note that model selection criteria, such as the Bayesian Information Criterion (BIC, (Schwartz 1978)) or the Akaike Information Criterion (AIC, (Akaike 1974)), can be used to determine the appropriate model order p .

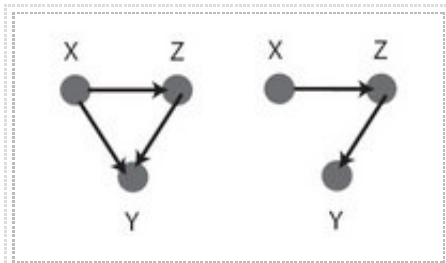


Figure 2: Two possible connectivities that cannot be distinguished by pairwise analysis. Adapted from Ding et al. (2006).

As mentioned in the previous section, G-causality can be readily extended to the n variable case, where $n > 2$, by estimating an n variable autoregressive model. In this case, X_2 G-causes X_1 if lagged observations of X_2 help predict X_1 when lagged observations of all other variables $X_3 \dots X_N$ are also taken into account. (Here, ' $X_3 \dots X_N$ ' correspond to the variables in the set W in the previous section; see also Boudjellaba et al. (1992) for an interpretation using autoregressive moving average (ARMA) models.) This multivariate extension, sometimes referred to as 'conditional' G-causality (Ding et al. 2006), is extremely useful because repeated pairwise analyses among multiple variables can sometimes give misleading results. For example, a

repeated bivariate analyses would be unable to disambiguate the two connectivity patterns in Figure 2. By contrast, a conditional/multivariate analysis would infer a causal connection from X to Y only if past information in X helped predict future Y above and beyond those signals mediated by Z . Another instance in which conditional G-causality is valuable is when a single source drives two outputs with different time delays. A bivariate analysis, but not a multivariate analysis, would falsely infer a causal connection from the output with the shorter delay to the output with the longer delay.

Application of the above formulation of G-causality makes two important assumptions about the data: (i) that it is covariance stationary (i.e., the mean and variance of each time series do not change over time), and (ii) that it can be adequately described by a linear model. The #Limitations and extensions section will describe recent extensions that attempt to overcome these limitations.

Spectral G-causality

By using Fourier methods it is possible to examine G-causality in the spectral domain (Geweke 1982; Kaminski et al. 2001). This can be very useful for neurophysiological signals, where frequency decompositions are often of interest. Intuitively, spectral G-causality from X_1 to X_2 measures the fraction of the total power at frequency f of X_1 that is contributed by X_2 . For examples of the advantages of working in the frequency domain see #Application in neuroscience.

For completeness, we give below the mathematical details of spectral G-causality. The Fourier transform of (1) gives

$$\begin{pmatrix} A_{11}(f) & A_{12}(f) \\ A_{21}(f) & A_{22}(f) \end{pmatrix} \begin{pmatrix} X_1(f) \\ X_2(f) \end{pmatrix} = \begin{pmatrix} E_1(f) \\ E_2(f) \end{pmatrix} \quad (2)$$

in which the components of A are

$$A_{lm}(f) = \delta_{lm} - \sum_{j=1}^p A_{lm}(j) e^{(-i2\pi f j)}$$

$$\delta_{lm} = 0 \quad (l = m)$$

$$\delta_{lm} = 1 \quad (l \neq m)$$

Rewriting Equation (2) as

$$\begin{pmatrix} X_1(f) \\ X_2(f) \end{pmatrix} = \begin{pmatrix} H_{11}(f) & H_{12}(f) \\ H_{21}(f) & H_{22}(f) \end{pmatrix} \begin{pmatrix} E_1(f) \\ E_2(f) \end{pmatrix}$$

we have

$$\begin{pmatrix} H_{11}(f) & H_{12}(f) \\ H_{21}(f) & H_{22}(f) \end{pmatrix} = \begin{pmatrix} A_{11}(f) & A_{12}(f) \\ A_{21}(f) & A_{22}(f) \end{pmatrix}^{-1}$$

where H is the transfer matrix. The spectral matrix S can now derived as

$$S(f) = \langle X(f)X^*(f) \rangle = \langle H(f)\Sigma H^*(f) \rangle$$

in which the asterisk denotes matrix transposition and complex conjugation, Σ is the covariance matrix of the residuals $E(t)$, and H is the transfer matrix. The spectral G-causality from j to i is then

$$I_{j \rightarrow i}(f) = -\ln \left(1 - \frac{\left(\Sigma_{jj} - \frac{\Sigma_{ij}^2}{\Sigma_{ii}} \right) |H_{ij}(f)|^2}{S_{ii}(f)} \right)$$

in which $S_{ii}(f)$ is the power spectrum of variable i at frequency f . (This analysis was adapted from (Brovelli et al. 2004; Kaminski et al. 2001)).

Recent work by Chen et al. (2006) indicates that application of Geweke's spectral G-causality to multivariate (>2) neurophysiological time series sometimes results in negative causality at certain frequencies, an outcome which evades physical interpretation. They have suggested a revised, conditional version of Geweke's measure which may overcome this problem by using a partition matrix method. Other variations of spectral G-causality are discussed by Breitung and Candelon (2006) and Hosoya (1991).

Two alternative measures which are closely related to spectral G-causality are partial directed coherence (Baccala & Sameshima 2001) and the directed transfer function (Kaminski et al. 2001; note these authors showed an equivalence between the directed transfer function and spectral G-causality). For comparative results among these methods see Baccala and Sameshima (2001), Gourevitch et al. (2006), and Pereda et al. (2005). Unlike the original time-domain formulation of G-causality, the statistical properties of these spectral measures have yet to be fully elucidated. This means that significance testing often relies on surrogate data, and the influence of signal pre-processing (e.g., smoothing, filtering) on measured causality remains unclear.

Limitations and extensions

Linearity

The original formulation of G-causality can only give information about linear features of signals. Extensions to nonlinear cases now exist, however these extensions can be more difficult to use in practice and their statistical properties are less well understood. In the approach of Freiwald et al. (1999) the globally nonlinear data is divided into a locally linear neighborhoods (see also Chen et al. 2004), whereas Ancona et al. (2004) used a radial basis function method to perform a global nonlinear regression.

Stationarity

The application of G-causality assumes that the analyzed signals are covariance stationary. Non-stationary data can be treated by using a windowing technique (Hesse et al. 2003) assuming that sufficiently short windows of a non-stationary signal are locally stationary. A related approach takes advantage of the trial-by-

Information theory

Schreiber (2000) introduced the concept of transfer entropy which is a version of mutual information operating on conditional probabilities. It is designed to detect the directed exchange of information between two variables, conditioned to common history and inputs. For a comparison of transfer entropy with other causal measures, including various implementations of G-causality, see (Lungarella et al. (in press)). An advantage of information theoretic measures, as compared to standard G-causality, is that they are sensitive to nonlinear signal properties. A limitation of transfer entropy, as compared to G-causality, is that it is currently restricted to bivariate situations. Also, information theoretic measures often require substantially more data than regression methods such as G-causality (Pereda et al., 2005).

Maximum likelihood models

For neural data consisting of spike trains, it is possible to use a combination of point process theory and maximum likelihood modeling to detect causal interactions within an ensemble of neurons (Chornoboy et al. 1988; Okatan et al. 2005).

Alternative frameworks

The above techniques (including G-causality) are *data driven* inasmuch as causal interactions are inferred directly from simultaneously recorded time series. A different *model driven* framework for analyzing causality is given by structural equation modeling (Kline 2005). In this approach, a causal model is first hypothesized, free parameters are then estimated, and only then is the fit of the model to the data assessed. While this method makes more efficient use of data than standard G-causality, it does so at the expense of constraining the repertoire of possible causal descriptions.

Another alternative framework for determining causality is to measure the effects of selective perturbations or lesions (Keinan et al. 2004; Tononi & Sporns 2003). While this method can in principle deliver unambiguous information about physical causal chains, perturbing or lesioning a system is not always possible or desirable and may disrupt the natural behavior of the system.

Concluding remarks

Time series analysis methods are becoming increasingly prominent in attempts to understand the relationships between network structure and network dynamics in neuroscience settings. Many linear and nonlinear methods now exist, based on a variety of techniques including regression modeling, information theory, and dynamical systems theory (Gourevitch et al., 2006; Pereda et al. 2005). G-causality provides one method that, in its most basic form (see #Mathematical formulation), is easy to implement and rests on a firm statistical foundation. More complex extensions to the frequency domain and to nonlinear data now exist and continue to be developed. However, in the analysis of neurophysiological signals it may be that simple, linear methods should be tried first before moving on to more complicated alternatives.

Resources

A detailed review of the theory and application of G-causality can be found in Ding et al. (2006).

James P. LeSage has provided a toolbox that includes G-causality analysis among a wide selection of econometric routines. This toolbox, designed for MATLAB (Mathworks, Natick, MA), can be downloaded from www.spatial-econometrics.com.

Anil K. Seth has developed a toolbox that focuses on time-domain G-causality and which includes methods for graph-theoretic analysis of G-causality interactions. This toolbox, also designed for MATLAB, can be downloaded from [Anil K Seth's website \(http://www.anilseth.com\)](http://www.anilseth.com).

Links to many other MATLAB resources for a variety of nonlinear time series analysis methods are provided in the Appendix of Pereda et al. (2005).

See also

transfer entropy, mutual information, structural equation modeling

External links

[Anil K Seth's website \(http://www.anilseth.com\)](http://www.anilseth.com).

References

- Akaike, H. 1974 A new look at the statistical model identification. *IEEE Trans. Autom. Control* 19, 716-723.
- Ancona, N., Marinazzo, D. & Stramaglia, S. 2004 Radial basis function approaches to nonlinear Granger causality of time series. *Physical Review E* 70, 056221.
- Baccala, L. A. & Sameshima, K. 2001 Partial directed coherence: a new concept in neural structure determination. *Biol Cybern* 84, 463-74.
- Bernasconi, C. & Konig, P. 1999 On the directionality of cortical interactions studied by structural analysis of electrophysiological recordings. *Biol Cybern* 81, 199-210.
- Boudjellaba, H., Dufour, J. & Roy, R. 1992 Testing causality between two vectors in multivariate autoregressive moving average models. *Journal of the American Statistical Association* 87, 1082-90.
- Breitung, J. & Candelon, B. 2006 Testing for short- and long-run causality: A frequency domain approach. *Journal of Econometrics* 132, 363-378.
- Brovelli, A., Ding, M., Ledberg, A., Chen, Y., Nakamura, R. & Bressler, S. L. 2004 Beta oscillations in a large-scale sensorimotor cortical network: directional influences revealed by Granger causality. *Proc Natl Acad Sci U S A* 101, 9849-54.
- Chen, Y., Bressler, S. L. & Ding, M. 2006 Frequency decomposition of conditional Granger causality and application to multivariate neural field potential data. *J Neurosci Methods* 150, 228-37.
- Chen, Y., Rangarajan, G., Feng, J. & Ding, M. 2004 Analyzing multiple nonlinear time series with extended Granger causality. *Physics Letters A* 324, 26-35.
- Chornoboy, E. S., Schramm, L. P. & Karr, A. F. 1988 Maximum likelihood identification of neural point process systems. *Biol Cybern* 59, 265-75.
- Ding, M., Chen, Y., & Bressler, S.L. 2006 Granger causality: Basic theory and application to neuroscience. In Schelter, S., Winterhalder, N., & Timmer, J. *Handbook of Time Series Analysis*. Wiley, Weinheim.
- Ding, M., Bressler, S. L., Yang, W. & Liang, H. 2000 Short-window spectral analysis of cortical event-

related potentials by adaptive multivariate autoregressive modeling: data preprocessing, model validation, and variability assessment. *Biological Cybernetics* 83, 35-45.

- Freiwald, W. A., Valdes, P., Bosch, J., Biscay, R., Jimenez, J. C., Rodriguez, L. M., Rodriguez, V., Kreiter, A. K. & Singer, W. 1999 Testing non-linearity and directedness of interactions between neural groups in the macaque inferotemporal cortex. *J Neurosci Methods* 94, 105-19.
- Geweke, J. 1982 Measurement of linear dependence and feedback between multiple time series. *Journal of the American Statistical Association* 77, 304-313.
- Gourevitch, B., Bouquin-Jeannes, R. L. & Faucon, G. 2006 Linear and nonlinear causality between signals: methods, examples and neurophysiological applications. *Biol Cybern* 95, 349-369.
- Granger, C. W. J. 1969 Investigating causal relations by econometric models and cross-spectral methods. *Econometrica* 37, 424-438.
- Granger, C. W. J. 1980 Testing for causality: A personal viewpoint. *Journal of Economic Dynamics and Control* 2, 329-352.
- Granger, C. W. J. 2001 *Essays in Econometrics: The Collected Papers of Clive W.J. Granger*. Cambridge: Cambridge University Press.
- Hesse, W., Moller, E., Arnold, M. & Schack, B. 2003 The use of time-variant EEG Granger causality for inspecting directed interdependencies of neural assemblies. *J Neurosci Methods* 124, 27-44.
- Hosoya, Y. 1991 The decomposition and measurement of the interdependence between second-order stationary processes. *Probability theory and related fields* 88, 429-444.
- Kaminski, M., Ding, M., Truccolo, W. A. & Bressler, S. L. 2001 Evaluating causal relations in neural systems: Granger causality, directed transfer function and statistical assessment of significance. *Biol Cybern* 85, 145-57.
- Keinan, A., Sandbank, B., Hilgetag, C. C., Meilijson, I. & Ruppin, E. 2004 Fair attribution of functional contribution in artificial and biological networks. *Neural Comput* 16, 1887-915.
- Kline, R. B. 2005 *Principles and Practices of Structural Equation Modelling*: The Guildford Press.
- Liang, H., Ding, M., Nakamura, R. & Bressler, S. L. 2000 Causal influences in primate cerebral cortex during visual pattern discrimination. *Neuroreport* 11, 2875-80.
- Lungarella, M., Ishiguro, K., Kuniyoshi, Y. & Otsu, N. (in press) Methods for quantifying the causal structure of bivariate time series. *International journal of bifurcation and chaos*.
- Okatan, M., Wilson, M. A. & Brown, E. N. 2005 Analyzing functional connectivity using a network likelihood model of ensemble neural spiking activity. *Neural Comput* 17, 1927-61.
- Pereda, E., Quiroga, R. Q. & Bhattacharya, J. 2005 Nonlinear multivariate analysis of neurophysiological signals. *Prog Neurobiol* 77, 1-37.
- Roebroeck, A., Formisano, E. & Goebel, R. 2005 Mapping directed influence over the brain using Granger causality and fMRI. *Neuroimage* 25, 230-42.
- Sato, J. R., Junior, E. A., Takahashi, D. Y., de Maria Felix, M., Brammer, M. J. & *Moretting, P. A. 2006 A method to produce evolving functional connectivity maps during the course of an fMRI experiment using wavelet-based time-varying Granger causality. *Neuroimage* 31, 187-96.
- Schreiber, T. 2000 Measuring information transfer. *Phys Rev Lett* 85, 461-4.
- Schwartz, G. 1978 Estimating the dimension of a model. *The Annals of Statistics* 5, 461-464.
- Seth, A. K. 2005 Causal connectivity analysis of evolved neural networks during behavior. *Network: Computation in Neural Systems* 16, 35-55.

- Seth, A. K. & Edelman, G. M. 2007 Distinguishing causal interactions in neural populations. *Neural Comput.* 19(4): 910-933
- Tononi, G. & Sporns, O. 2003 Measuring information integration. *BMC Neurosci* 4, 31.
- Wiener, N. 1956 The theory of prediction. In *Modern Mathematics for Engineers*, vol. 1 (ed. E. F. Beckenbach). New York: McGraw-Hill.

Internal references

- Yuri A. Kuznetsov (2007) Conjugate maps. *Scholarpedia*, 2(12):5420.
- James Meiss (2007) Dynamical systems. *Scholarpedia*, 2(2):1629.
- Paul L. Nunez and Ramesh Srinivasan (2007) Electroencephalogram. *Scholarpedia*, 2(2):1348.
- Rob Schreiber (2007) MATLAB. *Scholarpedia*, 2(7):2929.

See Also

Sponsored by: Eugene M. Izhikevich, Editor-in-Chief of Scholarpedia, the peer-reviewed open-access encyclopedia

Reviewed by (http://www.scholarpedia.org/w/index.php?title=Granger_causality&oldid=14741) :

Anonymous

Reviewed by (http://www.scholarpedia.org/w/index.php?title=Granger_causality&oldid=14741) : Dr. Clive Granger, Economics Department, University of California, San Diego, CA

Accepted on: 2007-07-06 23:15:45 GMT (http://www.scholarpedia.org/w/index.php?title=Granger_causality&oldid=15501)

Categories: Computational Neuroscience | Information Theory | Signal Analysis

"Granger causality" by Anil Seth is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License. Permissions beyond the scope of this license are described in the Terms of Use

