Granger Causality

Moo K. Chung
University of Wisconsin-Madison
mkchung@wisc.edu

1 Introduction

Causality is a fundamental concept in science, aiming to determine whether one event influences another. In time series analysis, **Granger causality**, introduced by Clive Granger in 1969, provides a statistical framework for assessing whether past values of one time series help predict another. Unlike traditional notions of causality in physics or philosophy, Granger causality is based on predictability if including past values of one time series improves the forecast of another. Granger causality has wide applications in economics, neuroscience, climate science, and engineering. In a *multivariate time series*, where multiple variables evolve over time, Granger causality extends naturally to examine interactions among multiple time-dependent signals. This is most often used causal model in brain network analysis. Granger causality is usually built on top of Vector Autoregressive (VAR) models.

2 Vector Autoregressive (VAR) Model

The **Vector Autoregressive (VAR)** model is a fundamental framework for modeling multivariate time series, where multiple variables evolve over time and potentially influence each other. Unlike univariate autoregressive models that describe the evolution of a single time series, the VAR model extends this concept to a system of interdependent time series. This makes it particularly useful for capturing dynamic relationships in fields such as economics, neuroscience, and climate science.

Consider a system of N time series denoted as

$$X(t) = [X_1(t), X_2(t), \dots, X_N(t)]^\top,$$

where $X_i(t)$ represents the time series of the i-th region at time t. The VAR model of order P, denoted as VAR(P), expresses each time series as a linear combination of its own past values and the past values of all other variables in the system. $X(t) \in \mathbb{R}^N$ is a time-varying vector. This model is formulated as

$$X(t) = \sum_{p=1}^{P} A_p X(t-p) + \varepsilon(t), \tag{1}$$

where A_p are $N \times N$ coefficient matrices that capture the influence of past values at lag p, and $\varepsilon(t)$ is an N-dimensional vector of error terms, assumed to be normally distributed with zero mean and a covariance matrix Σ, i.e.,

$$\varepsilon(t) \sim \mathcal{N}(0, \Sigma).$$

The number of lags P is a crucial hyperparameter that determines how far back in time the model considers past information. Expanding each component, we can write

$$X_i(t) = \sum_{p=1}^{P} \sum_{j=1}^{N} A_{ij}^{(p)} X_j(t-p) + \varepsilon_i(t), \tag{2}$$

where $A_{ij}^{(p)}$ represents the effect of the j-th variable at lag p on the i-th variable. Note $A_p = (A_{ij}^{(p)})$. If $A_{ij}^{(p)}$ is significantly different from zero, this indicates a directional relationship where past values of $X_j(t)$ contribute to predicting $X_i(t)$, which is the foundation for testing Granger causality.

The estimation of VAR model parameters is typically performed using least squares estimation:

$$\min_{A_1,\cdots,A_P} \sum_{t=P+1}^{T} \|X(t) - \sum_{p=1}^{P} A_p X(t-p)\|^2. \tag{3}$$

Instead of estimating the coefficients iteratively, which will compound errors, we can estimate the entire coefficient matrix simultaneously in a single step by reformulating the problem in matrix form. Let

$$A = [A_1, A_2, \ldots, A_P] \in \mathbb{R}^{N \times (NP)}, \tag{4}$$

which stacks all lag-specific coefficient matrices horizontally. Next, define the regressor matrix:

$$X_{\text{reg}}(t) = [X(t-1)^\top, X(t-2)^\top, \ldots, X(t-P)^\top]^\top \in \mathbb{R}^{(NP)\times 1}. \tag{5}$$

This allows the VAR model to be rewritten in a compact form as

$$X(t) = AX_{\text{reg}}(t) + \varepsilon(t), \quad t = P+1, P+2, \cdots T \tag{6}$$

where $\varepsilon(t)$ is the residual error term assumed to follow a multivariate normal distribution with zero mean and covariance Σ, i.e.,

$$\varepsilon(t) \sim \mathcal{N}(0, \Sigma). \tag{7}$$

Stacking the equations over all time points from $P+1$ to T, we define

$$Y = [X(P+1), X(P+2), \ldots, X(T)] \in \mathbb{R}^{N \times (T-P)}, \tag{8}$$

where Y represents the matrix of observed time series from $P+1$ to T. Similarly we define

$$\mathbf{X}_{\text{reg}} = [X_{\text{reg}}(P+1), X_{\text{reg}}(P+2), \ldots, X_{\text{reg}}(T)] \in \mathbb{R}^{(NP) \times (T-P)}, \tag{9}$$

Then we can write the VAR as a single matrix equation as

$$Y = A\mathbf{X}_{\text{reg}} + E, \tag{10}$$

where E is the residual error matrix

$$E = [\varepsilon(P+1), \varepsilon(P+2), \ldots, \varepsilon(T)] \in \mathbb{R}^{N \times (T-P)}. \tag{11}$$

To estimate A, we minimize the sum of squared residuals:

$$\widehat{A} = \arg\min_{A} \|Y - A\mathbf{X}_{\text{reg}}\|_F^2, \tag{12}$$

where $\|\cdot\|_F$ denotes the Frobenius norm. The least squares solution is given by

$$\widehat{A} = Y\mathbf{X}_{\text{reg}}^{\top} \left(\mathbf{X}_{\text{reg}}\mathbf{X}_{\text{reg}}^{\top}\right)^{-1}. \tag{13}$$

Once the coefficient matrix $\widehat{A}$ is obtained, the predicted values for all time points can be computed as

$$X_{\text{pred}} = \widehat{A}\mathbf{X}_{\text{reg}}. \tag{14}$$

This formulation ensures that all parameters are estimated simultaneously in a single step, minimizing the total squared error over all time points. The simultaneous estimation method provides a more stable and efficient approach compared to iterative forecasting methods, particularly when dealing with high-dimensional time series data.

The Vector Autoregressive (VAR) model requires past values of the time series $X(1), \ldots, X(P)$ to make predictions. The selection of the optimal lag order P is an essential step in model fitting. Choosing too small a value may result in missing important dependencies, whereas an excessively large P may introduce overfitting. The lag order is typically determined using information criteria such as the Akaike Information Criterion (AIC) or the Bayesian Information Criterion (BIC), which balance model complexity with goodness of fit.

Despite its flexibility and interpretability, the VAR model assumes stationarity and its statistical properties of the time series remain constant over time. Additionally, VAR models assume linear relationships, which may not fully capture nonlinear dependencies present in brain network data.

The VAR framework provides a powerful foundation for analyzing multivariate time series and forms the basis for testing Granger causality. By capturing interactions among multiple variables, it serves as an essential tool for investigating directional dependencies and understanding the underlying structure of complex dynamic systems.

3 Granger Causality

Granger causality provides a statistical framework for determining whether the past values of one time series improve the prediction of another within a multivariate time series setting. Unlike traditional notions of causality, which imply a mechanistic link, Granger causality is based purely on predictive ability: if including past values of a variable $X_j(t)$ significantly reduces the prediction error of another variable $X_i(t)$, then $X_j(t)$ is said to **Granger-cause** $X_i(t)$.

Given a multivariate system modeled by a vector autoregressive process of order P (VAR(P)):

$$X_i(t) = \sum_{p=1}^{P} \sum_{j=1}^{N} A_{ij}^{(p)} X_j(t-p) + \varepsilon_i(t), \tag{15}$$

where $A_{ij}^{(p)}$ are the autoregressive coefficients capturing the influence of variable j on variable i at lag p, and $\varepsilon_i(t)$ is a zero-mean white noise process with covariance Σ. To test whether $X_j(t)$ Granger-causes $X_i(t)$, we compare two models:

- **Full model:** The standard VAR(P) equation, where $X_i(t)$ depends on the past values of all variables, including $X_j(t-p)$, given by:

$$X_i(t) = \sum_{p=1}^{P} \sum_{j=1}^{N} A_{ij}^{(p)} X_j(t-p) + \varepsilon_i(t). \tag{16}$$

- **Restricted model:** A modified VAR(P) equation where the past values of $X_j(t)$ are excluded from the predictors:

$$X_i(t) = \sum_{p=1}^{P} \sum_{j \neq k} A_{ij}^{(p)} X_j(t-p) + \varepsilon_i'(t). \tag{17}$$

We then determine whether the exclusion of $X_j(t-p)$ significantly worsens the prediction of $X_i(t)$. If the residual error in the restricted model increases significantly compared to the full model, we conclude that $X_j(t)$ provides useful predictive information for $X_i(t)$ and therefore Granger-causes it. To statistically test this, we define the hypotheses as follows:

$$H_0 : A_{ij}^{(p)} = 0, \quad \text{for all } p \in \{1, \ldots, P\} \tag{18}$$

$$H_A : A_{ij}^{(p)} \neq 0, \quad \text{for at least one } p. \tag{19}$$

Under the null hypothesis H_0, the past values of $X_j(t)$ do not contribute to predicting $X_i(t)$, implying no Granger causality. If the null hypothesis is rejected, it suggests that past values of $X_j(t)$ improve the prediction of $X_i(t)$, supporting the presence of Granger causality. Thus, we use the F-statistic computed as

$$F = \frac{(\mathrm{RSS}_{\mathrm{restricted}} - \mathrm{RSS}_{\mathrm{full}})/P}{\mathrm{RSS}_{\mathrm{full}}/(T-k)} \sim f_{P,T-k} \tag{20}$$

where: $\mathrm{RSS}_{\mathrm{restricted}}$ is the residual sum of squares for the restricted model, $\mathrm{RSS}_{\mathrm{full}}$ is the residual sum of squares for the full model. k is the total number of parameters in the full model computed as

$$k = N \times P \times N = N^2 P.$$

Under H_0, the corresponding p-value is computed as

$$p = 1 - F_{P,T-k}(F). \tag{21}$$

with the cumulative distribution $F_{P,T-k}$.

For large N number of regions, the test results are typically presented in a **Granger causality matrix** G, where each entry G_{ij} represents the p-value of the Granger causality test from X_j to X_i, testing whether past values of $X_j(t)$ significantly improve the prediction of $X_i(t)$. Lower p-values indicate stronger evidence of Granger causality. A directed graph can be constructed from G, where an edge from X_j to X_i indicates a significant causal relationship. This allows a visually displaying causal relationship of one brain region to another. Thus, Granger causality is widely used in neuroscience to infer directional functional connectivity between brain regions from fMRI and EEG data. By applying the VAR model to time-lagged connectivity matrices, researchers can identify key interactions and how information propagates across brain networks. Similarly, in econometrics, Granger causality is used to analyze the predictive relationships between financial indicators, such as interest rates and stock prices.