Multivariate Linear and Non-Linear Causality Tests

Zhidong Bai\textsuperscript{a}, Wing-Keung Wong\textsuperscript{b} and Bingzhi Zhang\textsuperscript{c}

\textsuperscript{a} School of Mathematics and Statistics, Northeast Normal University, China; and Department of Statistics and Applied Probability, National University of Singapore, Singapore
\textsuperscript{b} Department of Economics, Hong Kong Baptist University, Hong Kong
\textsuperscript{c} Department of BioStatistics, Columbia University, USA

ARTICLE INFO

Keywords:
linear Granger causality,
nonlinear Granger causality,
U-statistics

JEL classification codes:
C01, C12, G10

ABSTRACT

The traditional linear Granger test has been widely used to examine the linear causality among several time series in bivariate settings as well as multivariate settings. Hiemstra and Jones (1994) develop a nonlinear Granger causality test in a bivariate setting to investigate the nonlinear causality between stock prices and trading volume. In this paper, we first discuss linear causality tests in multivariate settings and thereafter develop a non-linear causality test in multivariate settings.

1. Introduction

It is an important issue to detect the causal relation among several time series and it starts with two series, see, for example, Chiang, et al (2010) and Qiao, et al (2008, 2009) and the references therein for more discussion. To examine whether past information of one series could contribute to the prediction of another series, Granger causality test (Granger, 1969) is developed to examine whether lag terms of one variable significantly explain another variable in a vector autoregressive regression model.

Linear Granger causality test can be used to detect the causal relation between two time series. However, the linear Granger causality test does not perform well in detecting nonlinear causal relationships. To circumvent this limitation, Baek and Brock (1992) develop a nonlinear Granger causality test to examine the remaining nonlinear predictive power of a residual series of a variable on the residual of another variable obtaining from a linear model. Hiemstra and Jones (1994) have further modified the test which enables academics and practitioners to examine the bivariate nonlinear causality relationship between two series. Nevertheless, the multivariate causal relationships are important but it has not been well-studied, especially for nonlinear causality relationship. Thus, it is important to extend the Granger causality test to nonlinear causality test in the multivariate settings.

In this paper, we first discuss linear causality tests in multivariate settings and thereafter develop a non-linear causality test in multivariate settings. For any $n$
variables involved in the causality test, we discuss a $n$-equation vector autoregressive regression (VAR) model to conduct the linear Granger test, and test for the significance of relevant coefficients across equations using likelihood ratio test. If those coefficients are significantly different from zero, the linear causality relationship is identified. We then extend the nonlinear Granger test from bivariate setting to multivariate setting. We notice that the bivariate nonlinear Granger test is developed by mainly applying the properties of U-statistic developed by Denker and Keller (1983, 1986). Central limit theorem can be applied to the U-statistic whose arguments are strictly stationary, weakly dependent and satisfy mixing conditions of Denker and Keller (1983, 1986). When we extend the test to the multivariate settings, we find that the properties of the U-statistic for the bivariate settings could also be used in the development of our proposed test statistic in the multivariate settings, which is also a function of U-statistic.

The paper is organized as follows. We begin in next section by introducing definitions and notations and stating some basic properties for the linear and nonlinear Granger causal tests to test for linear and nonlinear Granger causal relationships between two series. In Section 3, we first discuss the linear Granger causality tests in the multivariate settings and thereafter develop the nonlinear Granger causality tests in the multivariate settings Section 4 gives a summary of our paper.

2. Bivariate Granger Causality Test

In this section, we will review the definitions of linear and nonlinear causality and discuss the linear and nonlinear Granger causality tests to identify the causality relationships between two variables.

2.1 Bivariate Linear Granger Causality Test

The linear Granger causality is conducted based on the following two-equation model:

Definition 2.1. In a two-equation model:

\[ x_t = a_1 + \sum_{i=1}^{p} \alpha_i x_{t-i} + \sum_{i=1}^{p} \beta_i y_{t-i} + \varepsilon_{1t}, \quad (1a) \]

and

\[ y_t = a_2 + \sum_{i=1}^{p} \gamma_i x_{t-i} + \sum_{i=1}^{p} \delta_i y_{t-i} + \varepsilon_{2t}, \quad (1b) \]

where all \( \{x_t\} \) and \( \{y_t\} \) are stationary variables, \( p \) is the optimal lag in the system, and \( \varepsilon_{1t} \) and \( \varepsilon_{2t} \) are the disturbances satisfying the regularity assumptions of the classical linear normal regression model. The variable \( \{y_t\} \) is said not to Granger cause \( \{x_t\} \) if \( i \beta_i = 0 \) in \((1a)\), for any \( i = 1, \ldots, p \). In other words, the past values of \( \{y_t\} \) do not provide any additional information on the performance of \( \{x_t\} \). Similarly, \( \{x_t\} \) does not Granger cause \( \{y_t\} \) if \( i \gamma_i = 0 \) in \((1b)\), for any \( i = 1, \ldots, p \).

It is well-known that one can test for linear causal relations between \( \{x_t\} \) and \( \{y_t\} \) by testing the following null hypotheses separately:

\[ H_0^1: \beta_1 = \ldots = \beta_p = 0, \quad \text{and} \]

\[ H_0^2: \gamma_1 = \ldots = \gamma_p = 0. \]

From testing these hypotheses, we have four possible testing results:

(1) If both Hypotheses \( H_0^1 \) and \( H_0^2 \) are accepted, there is no linear causal relationship between \( \{x_t\} \) and \( \{y_t\} \).

(2) If Hypothesis \( H_0^1 \) is accepted but Hypothesis \( H_0^2 \) is rejected, then there exists linear causality running unidirectionally from \( \{x_t\} \) to \( \{y_t\} \).
If Hypothesis $H_0^1$ is rejected but Hypothesis $H_0^2$ is accepted, then there exists linear causality running unidirectionally from $\{x_t\}$ to $\{y_t\}$.

If both Hypotheses $H_0^1$ and $H_0^2$ are rejected, then there exist feedback linear causal relationships between $\{x_t\}$ and $\{y_t\}$.

There are several statistics could be used to test the above hypotheses. One of the most commonly used statistics is the standard $F$-test. To test the hypothesis $H_0$: $\beta_1 = \ldots = \beta_p = 0$ in (1a), the sum of squares of the residuals from both the full regression, $SS_{FE}$, and the restricted regression, $SS_{RE}$, are computed in the equation (1a) and the $F$ test follows

$$ F = \frac{(SSR_R - SSR_F)/p}{SSR_F/(n-2p-1)}, \quad (1) $$

where $p$ is the optimal number of lag terms of $y_t$ in the regression equation on $x_t$ and $n$ is the number of observations. If $\{y_t\}$ does not Granger cause $\{x_t\}$, $F$ in (1) is distributed as $F_{(p,n-2p-1)}$.

For any given significance level $\alpha$, we reject the null hypothesis $H_0^1$ if $F$ exceeds the critical value $F_{(\alpha,p,n-2p-1)}$. Similarly, one could test for the second null hypothesis $H_0^2$: $\gamma_1 = \ldots = \gamma_p = 0$, and identify the linear causal relationship from $\{x_t\}$ to $\{y_t\}$.

### 2.2 Bivariate Nonlinear Causality Test

The general test for nonlinear Granger causality is first developed by Baek and Brock (1992) and, later on, modified by Hiemstra and Jones (1994). As the linear Granger test is inefficient in identifying any nonlinear causal relationship, to examine the nonlinear Granger causality relationship between a pair of series, say $\{x_t\}$ and $\{y_t\}$, one has to first apply the linear models in (1a) and (1b) to $\{x_t\}$ and $\{y_t\}$ for identifying their linear causal relationships and obtain their corresponding residuals, $\{\hat{e}_{1t}\}$ and $\{\hat{e}_{2t}\}$.

Thereafter, one has to apply a non-linear Granger causality test to the residual series, $\{\hat{e}_{1t}\}$ and $\{\hat{e}_{2t}\}$, of the two variables, $\{x_t\}$ and $\{y_t\}$, being examined to identify the remaining nonlinear causal relationships between their residuals.

We first state the definition of nonlinear Granger causality as follows:

**Definition 2.2.** For any two strictly stationary and weakly dependent series $\{X_t\}$ and $\{Y_t\}$, the $m$-length lead vector of $X_t$ is given by

$$ X_t^m \equiv (X_t, X_{t+1}, \ldots, X_{t+m-1}), \quad m = 1, 2, \ldots, t = 1, 2, \ldots $$

and $L_x$-length lag vector of $X_t$ is defined as

$$ X_t^{L_x} \equiv (X_{t-L_x}, X_{t-L_x-1}, \ldots, X_t), \quad L_x = 1, 2, \ldots, t = L_x + 1, L_x + 2, \ldots $$

The $m$-length lead vector, $Y_t^m$, and the $L_x$-length lag vector, $X_t^{L_x}$, of $Y_t$ can be defined similarly. Series $\{Y_t\}$ does not strictly Granger cause another series $\{X_t\}$ nonlinearly if and only if:

$$ \Pr(||X_t^m - X_t^m|| < \epsilon \| X_t^{L_x} - X_t^{L_x} \| < \epsilon, \| X_t^{L_x} - X_t^{L_x} \| < \epsilon) = \Pr(||X_t^m - X_t^m|| < \epsilon \| X_t^{L_x} - X_t^{L_x} \| < \epsilon), $$

where $\Pr(\cdot)$ denotes conditional probability and $||\cdot||$ denotes the maximum norm which is defined as

$$ ||X - Y|| = \max(|x_1 - y_1|, |x_2 - y_2|, \ldots, |x_n - y_n|), $$

for any two vectors $X = (x_1, \ldots, x_n)$ and $Y = (y_1, \ldots, y_n)$.

Under Definition 2.2, the non-linear Granger causality test statistic is given by

$$ \sqrt{n} \left( \frac{C_4(m+L_x,L_y,e,n) - C_4(m+L_x,L_y,e,n)}{C_2(L_x,L_y,e,n) - C_4(L_x,e,n)} \right) \quad (2) $$
Where
\[ C_1(m+L_x, L_y, e, n) = \frac{2}{n(n-1)} \sum_{t=1}^{n} I(x_{t-L_x}^{m+L_x}, x_{t-L_y}^{m+L_y}, e). \]
\[ C_2(L_x, L_y, e, n) = \frac{2}{n(n-1)} \sum_{t=1}^{n} I(x_{t-L_x}^{L_x}, x_{t-L_y}^{L_y}, e). \]
\[ C_3(L_x, L_y, e, n) = \frac{2}{n(n-1)} \sum_{t=1}^{n} I(x_{t-L_x}^{L_x}, x_{t-L_y}^{L_y}, e), \]
and
\[ I(x, y, e) = \begin{cases} 0, & \text{if } \|x - y\| > e \\ 1, & \text{if } \|x - y\| \leq e \end{cases}. \]

The test statistic, see Hiemstra and Jones (1994), possesses the following property:

Theorem 2.1. For given values of m, L_x, L_y and e > 0 defined in Definition 2.2, under the assumptions that \{X_t\}, \{Y_t\} are strictly stationary, weakly dependent, and satisfy the conditions stated in Denker and Keller (1983), if \{Y_t\} does not strictly Granger cause \{X_t\}, then the test statistic defined in (2) is distributed as \(N(0, \sigma^2(m, L_x, L_y, e))\) asymptotically, and the estimator of the variance \(\sigma^2(m, L_x, L_y, e)\) is given by
\[ \hat{\sigma}^2(m, L_x, L_y, e) = \hat{d}^T \cdot \hat{\Sigma} \cdot \hat{d}, \]
where
\[ \hat{d} = \left[ \begin{array}{c} \frac{1}{C_2(L_x, L_y, e, n)} \cdot C_1(m+L_x, L_y, e, n) \\ - \frac{1}{C_2(L_x, L_y, e, n)} \cdot C_3(m+L_x, L_y, e, n) \\ \vdots \\ - \frac{1}{C_4(L_x, e, n)} \cdot C_2^2(L_x, e, n) \end{array} \right]^T, \]
\[ \hat{\Sigma} = \sum_{k=1}^{K(n)} \omega_k(n) \]
and \(\hat{\Sigma}_{i,j}\) is a matrix containing elements
\[ \hat{\Sigma}_{i,j} = 4 \cdot \sum_{k=1}^{K(n)} \omega_k(n) \]
in which \(K(n) = [n^{1/4}]\), \([x]\) is the integer part of \(x\),
\[ \omega_k(n) = \begin{cases} 1, & \text{if } k = 1 \\ 2(1 - ((k-1) / K(n))), & \text{otherwise} \end{cases} \]
\[ \hat{A}_{i,t} = \frac{1}{n-1} \left( \sum_{s=0}^{T-m} I(x_{t-L_x}^{m+L_x}, x_{t-L_y}^{m+L_y}, e) \cdot I(y_{t-L_y}^{L_y}, y_{t-L_y}^{L_y}, e) \right) \]
\[ - C_1(m+L_x, L_y, e, n), \]
\[ \hat{A}_{2,t} = \frac{1}{n-1} \left( \sum_{s=0}^{T-m} I(x_{t-L_x}^{L_x}, x_{t-L_y}^{L_y}, e) \cdot I(y_{t-L_y}^{L_y}, y_{t-L_y}^{L_y}, e) \right) \]
\[ - C_2(L_x, L_y, e, n), \]
\[ \hat{A}_{3,t} = \frac{1}{n-1} \left( \sum_{s=0}^{T-m} I(x_{t-L_x}^{L_x}, x_{t-L_y}^{L_y}, e) \right) - C_3(m+L_x, L_y, e, n), \]
and
\[ t, s = \max(L_x, L_y) + 1, ..., T - m + 1 \]

3. Multivariate Granger Causality Test

In this section, we first discuss the linear Granger causality tests in the multivariate settings and, thereafter, develop the non-linear Granger causality tests from the bivariate settings to the multivariate settings.

3.1 Multivariate Linear Granger Causality Test

We first discuss the linear Granger causality test from the bivariate settings to the multivariate settings.

3.1.1 Vector Autoregressive Regression

The linear Granger test is applied in the vector autoregressive regression (VAR) scheme. For \(t = 1, ..., T\), the \(n\)-variable VAR model is represented as:
Multivariate Linear and Non-Linear Causality Tests

\[
\begin{pmatrix}
  y_{i1} \\
  y_{i2} \\
  \vdots \\
  y_{in}
\end{pmatrix} =
\begin{pmatrix}
  A_{00} \\
  A_{01} \\
  \vdots \\
  A_{0n}
\end{pmatrix} +
\begin{pmatrix}
  A_{11}(L) A_{12}(L) \cdots A_{1n}(L) \\
  A_{21}(L) A_{22}(L) \cdots A_{2n}(L) \\
  \vdots \\
  A_{n1}(L) A_{n2}(L) \cdots A_{nn}(L)
\end{pmatrix}
\begin{pmatrix}
  y_{i,1-1} \\
  y_{i,2-1} \\
  \vdots \\
  y_{i,n-1}
\end{pmatrix}
\]

\[
\begin{pmatrix}
  e_{i1} \\
  e_{i2} \\
  \vdots \\
  e_{in}
\end{pmatrix},
\]

(3)

where \((y_{i1}, \ldots, y_{in})\) is the vector of \(n\) stationary time series at time \(t\), \(L\) is the backward operation in which \(Lx_i = x_{i-1}\), \(A_{i0}\) are intercept parameters, \(A_{ij}(L)\) are polynomials in the lag operator \(L\) such that

\[A_{ij}(L) = a_{ij}(1) L + a_{ij}(2) L^2 + \cdots + a_{ij}(p) L^p\]

and \(e_i = (e_{i1}, \ldots, e_{in})'\) is the disturbance vector obeying the assumption of the classical linear normal regression model.

In practice, it is common to set all the equations in VAR to possess the same lag length for each variable. So a uniform order \(p\) will be chosen for all the lag polynomials \(A_{ij}(L)\) in the VAR model according to a certain criteria such as Akaike's information criterion (AIC) or Schwarz criterion (SC). Along with the Gauss-Markov assumptions satisfied for the error terms, ordinary least square estimation (OLSE) is appropriate to be used to estimate the model as it is consistent and efficient. However, long lag length for each variable will consume large number of degrees of freedom. For example, in the model stated in equation 3, there will be \(n(np + 1)\) coefficients including \(n\) intercept terms, \(n\) variances and \(n(n-1)/2\) covariances to be estimated. When the available sample size \(T\) is not large enough, including too many regressors will make the estimation inefficient, and thus, cause the test unreliable. To circumvent this problem, one could adopt a Near-VAR model and seemingly unrelated regressions estimation technique instead of applying OLSE to estimate the equations simultaneously. We skip the discussion of the Near-VAR model and seemingly unrelated regressions estimation in this paper and, for simplicity, we only use OLSE to estimate the parameters in the VAR model to identify the causality relationship among vectors of different time series.

3.1.2 Multiple Linear Granger Causality Hypothesis and Likelihood Ratio Test

To test the linear causality relationship between two vectors of different stationary time series, \(x_t = (x_{t1}, \ldots, x_{tn})'\) and \(y_t = (y_{t1}, \ldots, y_{tn})'\), where there are \(n1 + n2 = n\) series in total, one could construct the following \(n\)-equation VAR as follows:

\[
\begin{pmatrix}
  x_t \\
  y_t
\end{pmatrix} =
\begin{pmatrix}
  A_{x[n1]}(L) \\
  A_{y[n2]}(L)
\end{pmatrix} +
\begin{pmatrix}
  A_{xx}(L)[n1,n1] A_{xy}(L)[n1,n2] \\
  A_{yx}(L)[n2,n1] A_{yy}(L)[n2,n2]
\end{pmatrix}
\begin{pmatrix}
  x_{t-1} \\
  y_{t-1}
\end{pmatrix}
\]

\[
+ \begin{pmatrix}
  e_x \\
  e_y
\end{pmatrix},
\]

(4)

Where \(A_{x[n1]}\) and \(A_{y[n2]}\) are two vectors of intercept terms, and \(A_{xx}(L)[n1,n1]A_{xy}(L)[n1,n2]\) and \(A_{yx}(L)[n2,n1]A_{yy}(L)[n2,n2]\) are matrices of lag polynomials.

Similar to the bivariate case, there are four different situations for the existence of linear causality relationships between two vectors of time series \(x_t\) and \(y_t\) in (4):

(1) There exists a unidirectional causality from \(y_t\) to \(x_t\) if \(A_{xy}(L)\) is significantly different from zero \(^1\) and, at the same time, \(A_{yx}(L)\) is not significantly different from zero;

\(^1\) We said \(A_{xy}(L)\) is significantly different from zero if there exists any term in \(A_{xy}(L)\) which is significantly different from zero.
(2) there exists a unidirectional causality from \(x_t\) to \(y_t\) if \(A_{yx}(L)\) is significantly different from zero and, at the same time, \(A_{xy}(L)\) is not significantly different from zero;

(3) there exist feedback relations when both \(A_{xy}(L)\) and \(A_{yx}(L)\) are significantly different from zero;

(4) \(x_t\) and \(y_t\) are not rejected to be independent when both \(A_{xy}(L)\) and \(A_{yx}(L)\) is not significantly different from zero.

We note that one could consider one more situation as follows:

(5) \(x_t\) and \(y_t\) are rejected to be independent when either \(A_{xy}(L)\) and \(A_{yx}(L)\) is significantly different from zero. This is the same situation as either (1), (2) or (3) is true.

To test the above statements is equivalent to test the following null hypotheses:

(1) \(H_0^1: A_{xy}(L) = 0\),

(2) \(H_0^2: A_{yx}(L) = 0\), and

(3) both \(H_0^1\) and \(H_0^2: A_{xy}(L) = 0\) and \(A_{yx}(L) = 0\).

One may first obtain the residual covariance matrix \(\Sigma\) from the full model in (4) by using OLSE for each equation without imposing any restriction on the parameters, and compute the residual covariance matrix \(\Sigma_0\) from the restricted model in (4) by using OLSE for each equation with the restriction on the parameters imposed by the null hypothesis, \(H_0^1\), \(H_0^2\), or both \(H_0^1\) and \(H_0^2\).

Thereafter, besides using the F-test in (1), one could also use the likelihood ratio test in (5) introduced in this paper to identify the linear causality relationship for two variables in the bivariate settings.

### 3.1.3 ECM-VAR model

Consider \((Y_1, ..., Y_n)\) to be a vector of \(n\) non-stationary time series and if cointegration exists with residual vector \(vecm_t\). Let \(y_{it} = \Delta Y_i\) for \(i = 1, ..., n\) be the corresponding stationary differencing series. In this situation, one should not use the VAR model as stated in (3), one should impose the error-correction mechanism (ECM) on the VAR to test for Granger causality between these variables. The ECM-VAR framework is:

\[
\begin{aligned}
\begin{pmatrix}
    y_{1t} \\
    y_{2t} \\
    \vdots \\
    y_{nt}
\end{pmatrix}
&= \begin{pmatrix}
    A_{10} \\
    A_{20} \\
    \vdots \\
    A_{n0}
\end{pmatrix} + \begin{pmatrix}
    A_{11}(L) & A_{12}(L) & \cdots & A_{1n}(L) \\
    A_{21}(L) & A_{22}(L) & \cdots & A_{2n}(L) \\
    \vdots & \vdots & \ddots & \vdots \\
    A_{n1}(L) & A_{n2}(L) & \cdots & A_{nn}(L)
\end{pmatrix} \begin{pmatrix}
    y_{1t-1} \\
    y_{2t-1} \\
    \vdots \\
    y_{nt-1}
\end{pmatrix} \\
&+ \begin{pmatrix}
    \alpha_1 \\
    \alpha_2 \\
    \vdots \\
    \alpha_n
\end{pmatrix} \cdot vecm_{t-1} + \begin{pmatrix}
    e_{1t} \\
    e_{2t} \\
    \vdots \\
    e_{nt}
\end{pmatrix},
\end{aligned}
\]
where \( ecm_{t-1} \) is the error correction term. In particular, in this paper, we consider to test the causality relationship between two non-stationary vector time series, \( X = (X_{1,t}, \ldots, X_{n,t})' \) and \( Y = (Y_{1,t}, \ldots, Y_{n,t})' \). \( x_{it} = \Delta X_{it} \) and \( y_{it} = \Delta Y_{it} \) are the corresponding stationary differencing series, where there are \( n_1 + n_2 = n \) series in total. If \( X_i \) and \( Y_i \) are cointegrated with residual vector \( vecm_t \), then, instead of using the \( n \)-equation VAR in (4), one should adopt the \( n \)-equation VAR in the following:

\[
\begin{pmatrix}
x_i \\
y_i
\end{pmatrix} = \begin{pmatrix}
A_{ix}(L)_{n \times n_1} & A_{iy}(L)_{n \times n_1} \\
A_{yx}(L)_{n_2 \times n} & A_{yy}(L)_{n_2 \times n}
\end{pmatrix} \begin{pmatrix}
x_{t-1} \\
y_{t-1}
\end{pmatrix} + \begin{pmatrix}
\alpha_{ix}(n_1 \times 1) \\
\alpha_{iy}(n_2 \times 1)
\end{pmatrix} \cdot vecm_{t-1} + \begin{pmatrix}
e_{x} \\
e_{y}
\end{pmatrix},
\]

(7)

where \( \alpha_{ix}(n_1 \times 1) \) and \( \alpha_{iy}(n_2 \times 1) \) are two vectors of intercept terms, \( A_{ix}(L)_{n \times n_1} \) and \( A_{iy}(L)_{n \times n_1} \), \( A_{yx}(L)_{n_2 \times n} \) and \( A_{yy}(L)_{n_2 \times n} \) are matrices of lag polynomials, \( \alpha_{ix}(n_1 \times 1) \) and \( \alpha_{iy}(n_2 \times 1) \) are the coefficient vectors for the error correction term \( ecm_t \). Thereafter, one should test the null hypothesis \( H_0 : A_{iy}(L) = 0 \) or \( H_0 : A_{yx}(L) = 0 \) to identify strict causality relation using the LR test as discussed in Section 3.1.2.

3.2 Multivariate Nonlinear Causality Test

In this section, we will extend the nonlinear causality test for a bivariate setting developed by Hiemstra and Jones (1994) to a multivariate setting.

3.2.1 Multivariate Nonlinear Causality Hypothesis

As discussed in Section 2.2, to identify any nonlinear Granger causality relationship from any two series, say \( \{x_i\} \) and \( \{y_i\} \) in a bivariate setting, one has to first apply the linear models in (1a) and (1b) to \( \{x_i\} \) and \( \{y_i\} \) to identify their linear causal relationships and obtain their corresponding residuals, \( \{\hat{e}_{x_i}\} \) and \( \{\hat{e}_{y_i}\} \). Thereafter, one has to apply a non-linear Granger causality test to the residual series, \( \{\hat{e}_{x_i}\} \) and \( \{\hat{e}_{y_i}\} \), of the two variables being examined to identify the remaining nonlinear causal relationships between their residuals. This is also true if one would like to identify existence of any nonlinear Granger causality relations between two vectors of time series, say \( x_t = (x_{1,t}, \ldots, x_{n,t})' \) and \( y_t = (y_{1,t}, \ldots, y_{n,t})' \) in a multivariate setting. One has to apply the \( n \)-equation VAR model in (4) or (7) to the series to identify their linear causal relationships and obtain their corresponding residuals. Thereafter, one has to apply a non-linear Granger causality test to the residual series instead of the original time series. For simplicity, in this section we will denote \( X = (X_{1,t}, \ldots, X_{n,t})' \) and \( Y = (Y_{1,t}, \ldots, Y_{n,t})' \) to be the corresponding residuals of any two vectors of variables being examined.

We first define the lead vector and lag vector of a time series, say \( X_{it} \), similar to the terms defined in Definition 2.2 as follows. For \( X_{it} \), \( i = 1, \ldots, n_1 \), the \( m_{x_i} \) -length lead vector and the \( L_{x_i} \) -length lag vector of \( X_{it} \) are defined, respectively, as

\[
X_{it}^{m_{x_i}} \equiv (X_{i,t-m_{x_i}}, \ldots, X_{i,t-1}), m_{x_i} = 1, 2, \ldots, t = 1, 2, \ldots, L_{x_i} = 1, 2, \ldots, t = L_{x_i} + 1, L_{x_i} + 2, \ldots .
\]

We denote \( M_{x_i} = (m_{x_1}, \ldots, m_{x_n}) \), \( L_x = (L_{x_1}, \ldots, L_{x_n}) \), \( m_x = \max(m_{x_1}, \ldots, m_{x_n}) \), and \( l_x = \max(L_{x_1}, \ldots, L_{x_n}) \). The \( m_{x_i} \) -length lead vector, \( Y_{i,t-L_{x_i}}^{m_{x_i}} \), and the \( L_{y_i} \) -length lag vector, \( Y_{i,t-L_{y_i}}^{L_{y_i}} \), of \( Y_{it} \), \( M_y, L_y, m_y \), and \( l_y \) can be defined similarly.

Given \( m_{x_i}, m_y, L_x, L_y, e_i \), we define the following four events:
(1) \[ \| X_{i,i}^M - X_{i,i}^M \| < \epsilon, \text{for any } i = 1, \ldots, n] \]

\[ = \| X_{i,i}^L - X_{i,i}^L \| < \epsilon, \text{for any } i = 1, \ldots, n] \];

(2) \[ \| X_{i,i}^L - X_{i,i}^L \| < \epsilon, \text{for any } i = 1, \ldots, n] \];

\[ = \| X_{i,i}^L - X_{i,i}^L \| < \epsilon, \text{for any } i = 1, \ldots, n] \];

(3) \[ \| Y_{i,i}^M - Y_{i,i}^M \| < \epsilon, \text{for any } i = 1, \ldots, n] \];

\[ = \| Y_{i,i}^L - Y_{i,i}^L \| < \epsilon, \text{for any } i = 1, \ldots, n] \];

(4) \[ \| Y_{i,i}^L - Y_{i,i}^L \| < \epsilon, \text{for any } i = 1, \ldots, n] \];

where \( || \cdot || \) denotes the maximum norm defined in Definition 2.2.

The vector series \( \{Y_i\} \) is said not to strictly Granger cause another vector series \( \{X_i\} \) if:

\[ \Pr(\| X_{i,i}^M - X_{i,i}^M \| < \epsilon | \| X_{i,i}^L - X_{i,i}^L \| < \epsilon) \]

\[ \Pr(\| Y_{i,i}^M - Y_{i,i}^M \| < \epsilon | \| Y_{i,i}^L - Y_{i,i}^L \| < \epsilon) \]

where \( \Pr(\cdot | \cdot) \) denotes conditional probability.

### 3.2.2 Test Statistic and It's Asymptotic Distribution

Similar to the bivariate case, the test statistic for testing non-existence of nonlinear Granger causality can be obtained as follows:

\[ \sqrt{n} \left( \frac{C_1(M_x + L_x, e, n) - C_2(L_x, L_y, e, n)}{C_2(L_x, L_y, e, n)} - \frac{C_3(M_x + L_x, e, n)}{C_4(L_x, e, n)} \right) \tag{8} \]

Where

\[ C_1(M_x + L_x, L_y, e, n) = \frac{2}{n(n-1)} \sum_{i<s} \prod_{i=1}^n I(x_{i,i}^{L_{i,i}}, x_{i,i}^{L_{i,i}}, \epsilon) \]

\[ \prod_{i=1}^n I(x_{i,i}^{L_{i,i}}, x_{i,i}^{L_{i,i}}, \epsilon), \]

and

\[ C_2(L_x, L_y, e, n) = \frac{2}{n(n-1)} \sum_{i<s} \prod_{i=1}^n I(x_{i,i}^{L_{i,i}}, x_{i,i}^{L_{i,i}}, \epsilon), \]

\[ C_3(M_x + L_x, e, n) = \frac{2}{n(n-1)} \sum_{i<s} \prod_{i=1}^n I(x_{i,i}^{M_{i,i}}, x_{i,i}^{M_{i,i}}, \epsilon), \]

and

\[ t,s = \max(L_x, L_y) + 1, \ldots, T - m_x + 1, n = T + l - m_x - \max(L_x, L_y). \]

#### Theorem 3.1

To test the null hypothesis, \( H_0 \), that \( \{Y_{i,i}, \ldots, Y_{n,i}\} \) does not strictly Granger cause \( \{X_{i,i}, \ldots, X_{n,i}\} \) under the assumptions that the time series \( \{X_{i,i}, \ldots, X_{n,i}\} \) and \( \{Y_{i,i}, \ldots, Y_{n,i}\} \) are strictly stationary, weakly dependent, and satisfy the mixing conditions stated in Denker and Keller (1983), if the null hypothesis, \( H_0 \), is true, the test statistic defined in (8) is distributed as \( N(0, \sigma^2(M_x, L_x, L_y, e)) \).

When the test statistic in (8) is too far away from zero, we reject the null hypothesis. A consistent estimator of \( \sigma^2(M_x, L_x, L_y, Lz, e) \) follows:

\[ \widehat{\sigma^2}(M_x, L_x, L_y, Lz, e) = \left( \frac{n}{2} \right)^{-1} \cdot \hat{\Sigma} \cdot \hat{\Sigma}^T \cdot \hat{\Sigma} \]

in which each component \( \Sigma_{i,j} \), \( i, j = 1, \ldots, 4 \), of the covariance matrix \( \Sigma \) is given by:

\[ \Sigma_{i,j} = 4 \cdot \sum_{k=1}^4 \omega_k E(A_{i,k} \cdot A_{j,k+1}), \]

\[ \omega_k = \begin{cases} 1 & \text{if } k = 1 \\ 2 & \text{otherwise}, \end{cases} \]
Multivariate Linear and Non-Linear Causality Tests

\[ A_{1,j} = h_1 \left( x_{t-L_x}^{M_x+L_x}, y_{t-L_y}^{L_y}, e \right) - C_1(M_x + L_x, L_y, e), \]

\[ A_{2,j} = h_2 \left( x_{t-L_x}^{L_x}, y_{t-L_y}^{L_y}, e \right) - C_2(L_x, L_y, e), \]

\[ A_{3,j} = h_3 \left( x_{t-L_x}^{M_x+L_x}, e \right) - C_3(M_x + L_x, e), \]

\[ A_{4,j} = h_4 \left( x_{t-L_x}^{L_x}, e \right) - C_4(L_x, e), \]

where \( h_i \left( z_i \right), \ i = 1,...,4, \) is the conditional expectation of \( h_i \left( z_i \right) \) given the value of \( z_i \) as follows:

\[ h_1 \left( x_{t-L_x}^{M_x+L_x}, y_{t-L_y}^{L_y}, e \right) = E \left( h_i \left| x_{t-L_x}^{M_x+L_x}, y_{t-L_y}^{L_y} \right. \right), \]

\[ h_2 \left( x_{t-L_x}^{L_x}, y_{t-L_y}^{L_y}, e \right) = E \left( h_i \left| x_{t-L_x}^{L_x}, y_{t-L_y}^{L_y} \right. \right), \]

\[ h_3 \left( x_{t-L_x}^{M_x+L_x}, e \right) = E \left( h_i \left| x_{t-L_x}^{M_x+L_x} \right. \right), \]

\[ h_4 \left( x_{t-L_x}^{L_x}, e \right) = E \left( h_i \left| x_{t-L_x}^{L_x} \right. \right). \]

A consistent estimator of \( \Sigma_{i,j} \) elements is given by:

\[ \hat{\Sigma}_{ij} = 4 \sum_{k=1}^{K(n)} \frac{1}{2(n-k+1)} \sum_i \left[ h_i \left( n \right) \cdot h_i \left( n \right) \cdot h_i \left( n \right) \cdot h_i \left( n \right) \right], \]

\[ K(n) = \left[ n \right], \]

\[ \omega_k \left( n \right) = \begin{cases} 1, & \text{if } k = 1 \\ 2(1 - [(k-1)/K(n)]), & \text{otherwise} \end{cases}, \]

\[ \hat{A}_{1,j} = \frac{1}{n-1} \left( \prod_{i=1}^{n} I \left( x_{t-L_x}^{m_x+L_x}, y_{t-L_y}^{m_y+L_y}, e \right) \right) \]

\[ \prod_{i=1}^{n} I \left( y_{t-L_y}^{m_y+L_y}, y_{t-L_y}^{m_y+L_y}, e \right) \]

\[ - C_1(M_x + L_x, L_y, e), \]

\[ \hat{A}_{2,j} = \frac{1}{n-1} \left( \prod_{i=1}^{n} I \left( L_x, L_y, e \right) \right) \]

\[ \prod_{i=1}^{n} I \left( y_{t-L_y}^{m_y+L_y}, y_{t-L_y}^{m_y+L_y}, e \right) \]

\[ - C_2(L_x, L_y, e), \]

and

\[ t = max(L_x, L_y), ... n \text{ and } n = T - m_x \]

\[ max(L_x, L_y) + 1, \] and a consistent estimator of \( \nabla f(\theta) \) is:

\[ \nabla f(\theta) = \left[ \frac{1}{\theta_1}, \frac{1}{\theta_2}, \frac{1}{\theta_3}, \frac{1}{\theta_4} \right]^T \]

\[ = \left[ \frac{1}{C_1(L_x, L_y, e)} - C_1(M_x + L_x, L_y, e) \right]

\[ \left[ \frac{1}{C_2(L_x, L_y, e)} - C_2(L_x, L_y, e) \right]^T \]

4. Conclusion Remarks

In this paper, we first discuss linear causality tests in multivariate settings and thereafter develop a non-linear causality test in multivariate settings. However, there is a disadvantage of the Hiemstra-Jones test. For example, Diks and Panchenko (2005) point out that Hiemstra-Jones test might have an over-rejection bias on the null hypothesis of Granger non-causality. Their simulation results show that rejection probability will goes to one as the sample size increases. Diks and Panchenko (2006) address this problem by replacing the global test by an average of local conditional dependence measures. Their new test shows weaker evidence for volume causing returns than Hiemstra-Jones test does. Besides Hiemstra-Jones test, other forms of nonlinear causality test has also been developed. For example, Marinazzo, Pellicoro, and Stramaglia (2008) adopt theory of reproducing kernel Hilber spaces to develop nonlinear Granger causality test. And Diks and DeGoede (2001) develop an information theoretic test statistics for Granger causality. They use bootstrap methods instead of asymptotic distribution to calculate the significance of
the test statistics. Thus, further extension of this paper could include to develop multivariate settings for the more powerful linear and non-linear causality tests.

References


