The Thailand Econometrics Society, Vol. 2, No. 2 (January 2010), 159 - 168

Multivariate Linear and Non-Linear Causality Tests

Zhidong Bai [a], Wing-Keung Wong[b] and Bingzhi Zhang [c]

[a] School of Mathematics and Statistics, Northeast Normal University, China; and Department of Statistics and Applied Probability, National University of Singapore, Singapore

[b] Department of Economics, Hong Kong Baptist University, Hong Kong

[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 rading 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.

* Corresponding author.
E-mail addresses: awong@hkbu.edu.hk.

Acknowledgments
The second author would like to thank Professors Robert B. Miller and Howard E. Thompson for their continuous guidance and encouragement. This research is partially supported by Northeast Normal University, the National University of Singapore, Hong Kong Baptist University, and Columbia University.

Linear Ganger 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 non-linear 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 non-linear 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 ε_{1t} and ε_{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 $\beta_i = 0$ in (1a), for any i = 1,…, 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 $\gamma_i = 0$ in (1b), for any i = 1,…, 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:

$H_0^1 : \beta_1 = ... = \beta_p = 0$, and

$H_0^2 : \gamma_1 = ... = \gamma_p = 0$.

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 $\{y_t\}$.

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

(4) 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 $H_0^1: \beta_1 = ... = \beta_p = 0$ in (1a), the sum of squares of the residuals from both the full regression, SSE_F, and the restricted regression, SSE_R, are computed in the equation (1a) and the F test follows

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

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 α, 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 $H_0^2: \gamma_1 = ... = \gamma_p = 0$, 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 detecting 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, $\{\hat{\varepsilon}_{1t}\}$ and $\{\hat{\varepsilon}_{2t}\}$. Thereafter, one has to apply a non-linear Granger causality test to the residual series, $\{\hat{\varepsilon}_{1t}\}$ and $\{\hat{\varepsilon}_{2t}\}$, 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*

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

and L_x-length lag vector of X_t is defined as

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

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

$$\Pr(\| X_t^m - X_s^m \| < e \,\big|\, \| X_{t-L_x}^{L_x} - X_{s-L_x}^{L_x} \| < e, \| X_{t-L_y}^{L_y} - X_{s-L_y}^{L_y} \| < e)$$

$$= \Pr(\| X_t^m - X_s^m \| < e \,\big|\, \| X_{t-L_x}^{L_x} - X_{s-L_x}^{L_x} \| < e),$$

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

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

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

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

$$\sqrt{n}\left(\frac{C_1(m+L_x, L_y, e, n)}{C_2(L_x, L_y, e, n)} - \frac{C_3(m+L_x, e, n)}{C_4(L_x, e, n)} \right) \qquad (2)$$

Where

$$C_1(m+L_x,L_y,e,n) \equiv \frac{2}{n(n-1)}\sum\sum_{t<s} I(x_{t-L_x}^{m+L_x}, x_{s-L_x}^{m+L_x}, e) \cdot I(y_{t-L_y}^{L_y}, y_{s-L_y}^{L_y}, e),$$

$$C_2(L_x,L_y,e,n) \equiv \frac{2}{n(n-1)}\sum\sum_{t<s} I(x_{t-L_x}^{L_x}, x_{s-L_x}^{L_x}, e) \cdot I(y_{t-L_y}^{L_y}, y_{s-L_y}^{L_y}, e),$$

$$C_3(m+L_x,e,n) \equiv \frac{2}{n(n-1)}\sum\sum_{t<s} I(x_{t-L_x}^{m+L_x}, x_{s-L_x}^{m+L_x}, e),$$

$$C_4(L_x,e,n) \equiv \frac{2}{n(n-1)}\sum\sum_{t<s} I(x_{t-L_x}^{L_x}, x_{s-L_x}^{L_x}, e),$$

and

$$I(x,y,e) = \begin{cases} 0, & if \ \|x-y\| > e \\ 1, & if \ \|x-y\| \le e \end{cases}.$$

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*

$$\sigma^2(m,L_x,L_y,e) = \hat{d}^T \cdot \hat{\Sigma} \cdot \hat{d},$$

where

$$\hat{d} = \left[\frac{1}{C_2(L_x,L_y,e,n)}, -\frac{C_1(m+L_x,L_y,e,n)}{C_2^2(L_x,L_y,e,n)}, -\frac{1}{C_4(L_x,e,n)}, \frac{C_3(m+L_x,e,n)}{C_4^2(L_x,e,n)} \right]^T$$

and $\hat{\Sigma}$ is a matrix containing elements

$$\hat{\Sigma}_{i,j} = 4 \cdot \sum_{k=1}^{K(n)} \omega_k(n) \left[\frac{1}{2(n-k+1)} \sum_t (\hat{A}_{i,t}(n) \cdot \hat{A}_{j,t-k+1}(n) + \hat{A}_{i,t-k+1}(n) \cdot \hat{A}_{j,t}(n)) \right],$$

in which $K(n) = [n^{1/4}]$, $[x]$ is the integer part of x,

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

$$\hat{A}_{1,t} = \frac{1}{n-1}\left(\sum_{s \ne t} I(x_{t-L_x}^{m+L_x}, x_{i,s-L_x}^{m+L_x}, e) \cdot I(y_{t-L_y}^{L_y}, y_{s-L_y}^{L_y}, e) \right) - C_1(m+L_x,L_y,e,n),$$

$$\hat{A}_{2,t} = \frac{1}{n-1}\left(\sum_{s \ne t} I(x_{t-L_x}^{L_x}, x_{s-L_x}^{L_x}, e) \cdot I(y_{t-L_y}^{L_y}, y_{s-L_y}^{L_y}, e) \right) - C_2(L_x,L_y,e,n),$$

$$\hat{A}_{3,t} = \frac{1}{n-1}\left(\sum_{s \ne t} I(x_{t-L_x}^{m+L_x}, x_{s-L_x}^{m+L_x}, e) \right) - C_3(m+L_x,e,n),$$

$$\hat{A}_{1,t} = \frac{1}{n-1}\left(\sum_{s \ne t} I(x_{t-L_x}^{L_x}, x_{s-L_x}^{L_x}, e) \right) - C_4(L_x,e,n),$$

and

$$t,s = \max(L_x,L_y)+1,\ldots,T-m+1$$

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 $t = 1,\ldots, T$, the n-variable VAR model is represented as:

$$\begin{pmatrix} y_{1t} \\ y_{2t} \\ \vdots \\ y_{nt} \end{pmatrix} = \begin{pmatrix} A_{10} \\ A_{20} \\ \vdots \\ A_{n0} \end{pmatrix} + \begin{pmatrix} A_{11}(L) & A_{21}(L) & \dots & A_{1n}(L) \\ A_{21}(L) & A_{22}(L) & \dots & A_{2n}(L) \\ \vdots & \vdots & \ddots & \vdots \\ A_{n1}(L) & A_{n2}(L) & \dots & A_{nn}(L) \end{pmatrix} \begin{pmatrix} y_{1,t-1} \\ y_{2,t-1} \\ \vdots \\ y_{n,t-1} \end{pmatrix}$$

$$+ \begin{pmatrix} e_{1t} \\ e_{2t} \\ \vdots \\ e_{nt} \end{pmatrix}, \qquad (3)$$

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

$$A_{ij}(L) = a_{ij}(1)L + a_{ij}(2)L^2 + ... + a_{ij}(p)L^p$$

and $e_t = (e_{1t}, ..., e_{nt})'$ 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, $x_t = (x_{1,t}, ..., x_{n1,t})'$ and $y_t = (y_{1,t}, ..., y_{n2,t})'$, where there are $n_1 + n_2 = n$ series in total, one could construct the following n-equation VAR as follows:

$$\begin{pmatrix} x_t \\ y_t \end{pmatrix} = \begin{pmatrix} A_{x[n_1\times 1]} \\ A_{y[n_2\times 1]} \end{pmatrix} + \begin{pmatrix} A_{xx}(L)_{[n_1\times n_1]} & A_{xy}(L)_{[n_1\times n_2]} \\ A_{yx}(L)_{[n_2\times n_1]} & A_{yy}(L)_{[n_2\times n_2]} \end{pmatrix} \begin{pmatrix} x_{t-1} \\ y_{t-1} \end{pmatrix}$$

$$+ \begin{pmatrix} e_x \\ e_y \end{pmatrix}, \qquad (4)$$

Where $A_{x[n_1\times 1]}$ and $A_{y[n_2\times 1]}$ are two vectors of intercept terms, and $A_{xx}(L)_{[n_1\times n_1]}$, $A_{xy}(L)_{[n_1\times n_2]}$, $A_{yx}(L)_{[n_2\times n_1]}$, and $A_{yy}(L)_{[n_2\times n_2]}$ 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 Σ 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 Σ_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 use a similar approach as in Sims (1980) to obtain the following likelihood ratio statistic:

$$(T-c)(\log|\Sigma_0| - \log|\Sigma|) \qquad (5)$$

where T is the number of usable observations, c is the number of parameters estimated in each equation of the unrestricted system, and $\log|\Sigma_0|$ and $\log|\Sigma|$ are the natural logarithms of the determinants of restricted and unrestricted residual covariance matrices, respectively. When the null hypothesis is true, this test statistic has a asymptotic χ^2 distribution with the degree of freedom equal to the number of restrictions on the coefficients in the system. For example, when we test $H_0 : A_{xy}(L) = 0$, one should let c equal to $np + 1$, and there are $n_2 \times p$ restrictions on the coefficients in the first n_1 equations of the model. Hence, the corresponding test statistic $(T-(np+1))(\log|\Sigma_0| - \log|\Sigma|)$ asymptotically follows χ^2 with $n_1 \times n_2 \times p$ degrees of freedom. The conventional bivariate causality test is an special case of the multivariate causality test when $n_1 = n_2 = 1$. Besides using the F test stated 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 $(Y_1, \ldots, Y_{nt})$ to be a vector of n non-stationary time series and if cointegration exists with residual vector $vecm_t$. Let $y_{it} = \Delta Y_{it}$ for $i = 1, \ldots, n$ 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:

$$\begin{pmatrix} y_{1t} \\ y_{2t} \\ \vdots \\ y_{nt} \end{pmatrix} = \begin{pmatrix} A_{10} \\ A_{20} \\ \vdots \\ A_{n0} \end{pmatrix} + \begin{pmatrix} A_{11}(L) & A_{21}(L) & \ldots & A_{1n}(L) \\ A_{21}(L) & A_{22}(L) & \ldots & A_{2n}(L) \\ \vdots & \vdots & \ddots & \vdots \\ A_{n1}(L) & A_{n2}(L) & \ldots & A_{nn}(L) \end{pmatrix} \begin{pmatrix} y_{1,t-1} \\ y_{2,t-1} \\ \vdots \\ y_{n,t-1} \end{pmatrix}$$

$$+ \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_{+n} \end{pmatrix} \cdot ecm_{t-1} + \begin{pmatrix} e_{1t} \\ e_{2t} \\ \vdots \\ e_{nt} \end{pmatrix}, \qquad (6)$$

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, $X_t = (X_{1,t}, ..., X_{n1,t})'$ and $Y_t = (Y_{1,t}, ..., Y_{n2,t})'$ and $x_{it} = \Delta X_{it}$ 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_t and Y_t 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:

$$\begin{pmatrix} x_t \\ y_t \end{pmatrix} = \begin{pmatrix} A_{x[n_1\times 1]} \\ A_{y[n_2\times 1]} \end{pmatrix} + \begin{pmatrix} A_{xx}(L)_{[n_1\times n_1]} & A_{xy}(L)_{[n_1\times n_2]} \\ A_{yx}(L)_{[n_2\times n_1]} & A_{yy}(L)_{[n_2\times n_2]} \end{pmatrix} \begin{pmatrix} x_{t-1} \\ y_{t-1} \end{pmatrix} + \begin{pmatrix} \alpha_{x[n_1\times 1]} \\ \alpha_{y[n_2\times 1]} \end{pmatrix} \cdot ecm_{t-1} + \begin{pmatrix} e_x \\ e_y \end{pmatrix}, \qquad (7)$$

where $\alpha_{x[n_1\times 1]}$ and $\alpha_{y[n_2\times 1]}$ are two vectors of intercept terms, $A_{xx}(L)_{[n_1\times n_1]}$, $A_{xy}(L)_{[n_1\times n_2]}$, $A_{yx}(L)_{[n_2\times n_1]}$, $A_{yy}(L)_{[n_2\times n_2]}$ are matrices of lag polynomials, $\alpha_{x[n_1\times 1]}$ and $\alpha_{y[n_2\times 1]}$ are the coefficient vectors for the error correction term $ecmt¡1$. Thereafter, one should test the null hypothesis $H_0 : A_{xy}(L) = 0$ 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 mulitvariate 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 $\{x_t\}$ and $\{y_t\}$ in a bivariate setting, one has to first apply the linear models in (1a) and (1b) to $\{x_t\}$ and $\{y_t\}$ to identify their linear causal relationships and obtain their corresponding residuals, $\{\hat{\varepsilon}_{1t}\}$ and $\{\hat{\varepsilon}_{2t}\}$. Thereafter, one has to apply a non-linear Granger causality test to the residual series, $\{\hat{\varepsilon}_{1t}\}$ and $\{\hat{\varepsilon}_{2t}\}$, 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 $x_t = (x_{1,t}, ..., x_{n1,t})'$ and $y_t = (y_{1,t}, ..., y_{n2,t})'$ 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 $X_t = (X_{1,t}, ..., X_{n1,t})'$ and $Y_t = (Y_{1,t}, ..., Y_{n2,t})'$ 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 $X_{i,t}$, similar to the terms defined in Definition 2.2 as follows. For $X_{i,t}$, $i = 1, ..., n_1$, the m_{x_i}-length lead vector and the L_{x_i}-length lag vector of $X_{i,t}$ are defined, respectively, as

$$X_{i,t}^{m_{x_i}} \equiv (X_{i,t}, X_{i,t+1}, ..., X_{i,t+m_{x_i}-1}), m_{x_i} = 1, 2, ..., t = 1, 2, ...,$$

$$X_{i,t-L_{x_i}}^{L_{x_i}} \equiv (X_{i,t-L_{x_i}}, X_{i,t-L_{x_i}+1}, ..., X_{i,t-1}),$$
$$L_{x_i} = 1, 2, ..., t = L_{x_i} + 1, L_{x_i} + 2, ... \, .$$

We denote $M_x = (m_{x_1}, ..., m_{x_{n_1}})$, $L_x = (L_{x_1}, ..., L_{x_{n_1}})$, $m_x = \max(m_{x_1}, ..., m_{x_{n_1}})$, and $l_x = \max(L_{x_1}, ..., L_{x_{n_1}})$,. The m_{y_i}-length lead vector, $Y_{i,t}^{m_{y_i}}$, and the L_{y_i}-length lag vector, $Y_{i,t-L_{y_i}}^{L_{y_i}}$, of $Y_{i,t}$, M_y, L_y, m_y, and l_y can be defined similarly.

Given m_x, m_y, L_x, L_y, e, we define the following four events:

(1) $\left\{\| X_t^{M_x} - X_s^{M_x} \| < e\right\}$

$\equiv\left\{\| X_{i,t}^{M_{x1}} - X_{i,s}^{M_{x1}} \| < e, for \quad any \quad i=1,...,n_1\right\};$

(2) $\left\{\| X_{t-L_x}^{L_x} - X_{s-L_x}^{L_x} \| < e\right\}$

$\equiv\left\{\| X_{i,t-L_{xi}}^{L_{xi}} - X_{i,s-L_{xi}}^{L_{xi}} \| < e, for \quad any \quad i=1,...,n_1\right\};$

(3) $\left\{\| Y_t^{M_y} - Y_s^{M_y} \| < e\right\}$

$\equiv\left\{\| Y_{i,t}^{M_{yi}} - Y_{i,s}^{M_{yi}} \| < e, for \quad any \quad i=1,...,n_2\right\}; and$

(4) $\left\{\| Y_{t-L_y}^{L_y} - Y_{s-L_y}^{L_y} \| < e\right\}$

$\equiv\left\{\| Y_{i,t-L_{yi}}^{L_{yi}} - Y_{i,s-L_{yi}}^{L_{yi}} \| < e, for \quad any \quad i=1,...,n_2\right\},$

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

The vector series $\{Y_t\}$ is said not to strictly Granger cause another vector series $\{X_t\}$ if:

$$\Pr\left(\| X_t^{M_x} - X_s^{M_x} \| < e \,\middle|\, \| X_{t-L_x}^{L_x} - X_{s-L_x}^{L_x} \| < e, \| Y_{t-L_y}^{L_y} - Y_{s-L_y}^{L_y} \| < e\right)$$
$$= \Pr\left(\| X_t^{M_x} - X_s^{M_x} \| < e \,\middle|\, \| X_{t-L_x}^{L_x} - X_{s-L_x}^{L_x} \| < e\right),$$

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:

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

Where

$$C_1(M_x+L_x,L_y,e,n)$$
$$\equiv \frac{2}{n(n-1)}\sum\sum_{t<s}\prod_{i=1}^{n_1} I(x_{i,t-L_{xi}}^{m_{xi}+L_{xi}}, x_{i,s-L_x}^{m_{xi}+L_{xi}}, e)$$
$$\cdot\prod_{i=1}^{n_2} I(y_{i,t-L_{yi}}^{L_{yi}}, y_{i,s-L_{yi}}^{L_{yi}}, e),$$

$$C_2(L_x,L_y,e,n)$$
$$\equiv \frac{2}{n(n-1)}\sum\sum_{t<s}\prod_{i=1}^{n_1} I(x_{i,t-L_{xi}}^{L_{xi}}, x_{i,s-L_x}^{L_{xi}}, e)$$
$$\cdot\prod_{i=1}^{n_2} I(y_{i,t-L_{yi}}^{L_{yi}}, y_{i,s-L_y}^{L_{yi}}, e),$$

$$C_3(M_x+L_x,e,n)$$
$$\equiv \frac{2}{n(n-1)}\sum\sum_{t<s}\prod_{i=1}^{n_1} I(x_{i,t-L_x}^{m_{xi}+L_{xi}}, x_{i,s-L_{xi}}^{m_{xi}+L_{xi}}, e),$$

$$C_4(L_x,e,n) \equiv \frac{2}{n(n-1)}\sum\sum_{t<s}\prod_{i=1}^{n_1} I(x_{i,t-L_{xi}}^{L_{xi}}, x_{i,s-L_{xi}}^{L_{xi}}, e),$$

and

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

Theorem 3.1. *To test the null hypothesis, H_0, that $\left\{Y_{1,t},...,Y_{n_2,t}\right\}$ does not strictly Granger cause $\left\{X_{1,t},...,X_{n_1,t}\right\}$ under the assumptions that the time series $\left\{X_{1,t},...,X_{n_1,t}\right\}$ and $\left\{Y_{1,t},...,Y_{n_2,t}\right\}$ 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 statisticin (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:*

$$\hat{\sigma}^2(M_x,L_x,L_y,Lz,e) = \widehat{\nabla f(\theta)}^T \cdot \hat{\Sigma} \cdot \widehat{\nabla f(\theta)}$$

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

$$\Sigma_{i,j} = 4\cdot\sum_{k\ge1}\omega_k E(A_{i,t}\cdot A_{j,t+k-1}),$$

$$\omega_k = \begin{cases} 1 & if \quad k=1 \\ 2, & otherwise \end{cases},$$

$$A_{1,t} = h_{11}\left(x_{t-L_x}^{M_x+L_x}, y_{t-L_y}^{L_y}, e\right) - C_1(M_x + L_x, L_y, e),$$

$$A_{2,t} = h_{12}\left(x_{t-L_x}^{L_x}, y_{t-L_y}^{L_y}, e\right) - C_2(L_x, L_y, e),$$
$$A_{3,t} = h_{13}\left(x_{t-L_x}^{M_x+L_x}, e\right) - C_3(M_x + L_x, e), and$$
$$A_{4,t} = h_{14}\left(x_{t-L_x}^{L_x}, e\right) - C_4(L_x, e),$$

where $h_{1i}\left(z_t\right),\ i = 1,...,4,$ *is the conditional expectation of* $h_i\left(z_t, z_s\right)$ *given the value of* z_t *as follows:*

$$h_{11}\left(x_{t-L_x}^{M_x+L_x}, y_{t-L_y}^{L_y}, e\right) = E\left(h_1 \middle| x_{t-L_x}^{M_x+L_x}, y_{t-L_y}^{L_y}\right),$$
$$h_{12}\left(x_{t-L_x}^{L_x}, y_{t-L_y}^{L_y}, e\right) = E\left(h_2 \middle| x_{t-L_x}^{L_x}, y_{t-L_y}^{L_y}\right),$$
$$h_{13}\left(x_{t-L_x}^{M_x+L_x}, e\right) = E\left(h_3 \middle| x_{t-L_x}^{M_x+L_x}\right), and$$
$$h_{14}\left(x_{t-L_x}^{L_x}, e\right) = E\left(h_4 \middle| x_{t-L_x}^{L_x}\right).$$

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

$$\hat{\Sigma}_{i,j} = 4 \cdot \sum_{k=1}^{K(n)} \omega_k(n)$$
$$\times \left[\frac{1}{2(n-k+1)} \sum_t \left(\hat{A}_{i,t}(n) \cdot \hat{A}_{j,t-k+1}(n) + \hat{A}_{i,t-k+1}(n) \cdot \hat{A}_{j,t}(n) \right) \right],$$

$$K(n) = [n^{1/4}],$$
$$\omega_k(n) = \begin{cases} 1 & if \quad k = 1 \\ 2(1 - [(k-1)/K(n)]), & otherwise \end{cases},$$

$$\hat{A}_{1,t} = \frac{1}{n-1}\left(\sum_{s \neq t} \prod_{i=1}^{n_1} I(x_{i,t-L_{x_i}}^{m_{x_i}+L_{x_i}}, x_{i,s-L_x}^{m_{x_i}+L_{x_i}}, e) \right.$$
$$\left. \cdot \prod_{i=1}^{n_2} I(y_{i,t-L_{y_i}}^{L_{y_i}}, y_{i,s-L_{y_i}}^{L_{y_i}}, e) \right)$$
$$-C_1(M_x + L_x, L_y, e, n),$$

$$\hat{A}_{2,t} = \frac{1}{n-1}\left(\sum_{s \neq t} \prod_{i=1}^{n_1} I(x_{i,t-L_{x_i}}^{L_{x_i}}, x_{i,s-L_x}^{L_{x_i}}, e) \right.$$
$$\left. \cdot \prod_{i=1}^{n_2} I(y_{i,t-L_{y_i}}^{L_{y_i}}, y_{i,s-L_y}^{L_{y_i}}, e) \right)$$
$$-C_2(L_x, L_y, e, n),$$

$$\hat{A}_{3,t} = \frac{1}{n-1}\left(\sum_{s \neq t} \prod_{i=1}^{n_1} I(x_{i,t-L_x}^{m_{x_i}+L_{x_i}}, x_{i,s-L_{x_i}}^{m_{x_i}+L_{x_i}}, e) \right) - C_3(m + L_x, e, n),$$
$$\hat{A}_{4,t} = \frac{1}{n-1}\left(\sum_{s \neq t} \prod_{i=1}^{n_1} I(x_{i,t-L_{x_i}}^{L_{x_i}}, x_{i,s-L_{x_i}}^{L_{x_i}}, e) \right) - C_4(L_x, e, n),$$

and

t,s=max(L_x,L_y),...,n and n=T-m_x-max(L_x,L_y)+1, and a consistent estimator of $\nabla f(\theta)$ *is:*

$$\widehat{\nabla f(\theta)} = \left[\frac{1}{\hat{\theta}_2}, -\frac{\hat{\theta}_1}{\hat{\theta}_2^2}, -\frac{1}{\hat{\theta}_4}, \frac{\hat{\theta}_3}{\hat{\theta}_4^2} \right]^T$$
$$= \left[\frac{1}{C_2(L_x, L_y, e, n)}, -\frac{C_1(m + L_x, L_y, e, n)}{C_2^2(L_x, L_y, e, n)}, \right.$$
$$\left. -\frac{1}{C_4(L_x, e, n)}, \frac{C_3(M_x + L_x, e, n)}{C_4^2(L_x, e, n)} \right]^T$$

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

Baek, E.G., and Brock, W.A., 1992, A General Test for Nonlinear Granger Causality : Bivariate Model. working paper, Korea Development Institut, University of Wisconsin-Madison.

Chiang, T.C., Qiao, Z., Wong, W.-K. 2010, New Evidence on the Relation between Return Volatility and Trading Volume. Journal of Forecasting, (forthcoming).

Denker, M., and Keller, G., 1983, On U-statistics and v. Mises' Statistics for weakly Dependent Processes, Zeitschrift fÄur Wahrscheinlichkeitstheorie und verwandte Gebiete 64, 505-522.

Denker, M., and Keller, G., 1986, Rigorous statistical procedures for data from dynamical system. Journal of Statistical Physics 44, 67-93.

Diks, C., and DeGoede, J., 2001, A general nonparametric bootstrap test for Granger causality. in: Global Analysis of Dynamical Systems, eds. H.W. Broer, B. Krauskopf and G. Vegter, 391-403, IoP Publishing, Bristol.

Diks, C. and Panchenko, V., 2005, A note on the Hiemstra-Jones test for Granger non-causality. Studies in Nonlinear Dynamics and Econometrics 9(2), 1-7.

Diks, C., and Panchenko, V., 2006, A new statistic and practical guidelines for nonparametric Granger causality testing. Journal of Economic Dynamics and Control 30(9-10), 1647-1669.

Granger, C.W.J., 1969, Investigating causal relations by econometric models and cross-spectral methods. Econometrica 37 (3), 424-438.

Hiemstra, C., and Jones, J.D., 1994, Testing for Linear and Nonlinear Granger Causality in the Stock Proce-Volume Relation. Journal of Finance 49(5), 1639-1664.

Kowalski, J., and Tu, X.M., 2007, Modern Applied U-Statistics. John Wiley & Sons, New York.

Marinazzo, D., Pellicoro, M., and Stramaglia, S., 2008, Kernel Method for Nonlinear Granger Causality. Physical Review Letters 100(14), Article 144103, 1-4.

Newey, W.K., and West, K.D., 1987, A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Constistent Covariance Matrix. Econometrica 55(3), 703-708.

Qiao, Z., Li, Y., Wong, W.-K., 2008, Policy Change and Lead-lag relations among China's Segmented Stock Markets. Journal of Multinational Financial Management 18, 276-289.

Qiao, Z., McAleer, M., Wong, W.-K., 2009, Linear and Nonlinear Causality of Consumption Growth and Consumer Attitudes. Economics Letters 102(3), 161-164.

Serfling, R., 1980, Approximation Theorems of Mathematical Statistics John Wiley & Sons, New York.

Sims, C.A., 1980, Macroeconomics and Reality. Econometrica 48, 1-48.