Notes On Testing Causality

Jin-Lung Lin
Institute of Economics, Academia Sinica
Department of Economics, National Chengchi University

March 8, 2007

Outline

Introduction

- Causal analysis is *very* important and yet, *complicated*
- Cannot perform controlled experiments to single out the impact of any particular variable.

Two most difficult challenges

1. Correlation does not imply causality. Distinguishing between these two is by no means an easy task.
2. There always exist the possibility of ignored common factors. The causal relationship among variables might disappear when the previously ignored common causes are considered.

Partial soultions:

- Granger causality: forecastability
- Directed graph theory

Two assumptions

1. The future cannot cause the past. The past causes the present or future. (How about expectation?)
2. A cause contains unique information about an effect not available elsewhere.

Definition

X_t is said not to Granger-cause Y_t if for all $h > 0$

$$F(Y_{t+h}|\Omega_t) = F(Y_{t+h}|\Omega_t - X_t)$$

where F denotes the conditional distribution and $\Omega_t - X_t$ is all the information in the universe except series X_t. In plain words, X_t is said not to Granger-cause Y_{t+h} if X cannot help predict future Y.

Remarks:

- The whole distribution F is generally difficult to handle empirically and we turn to conditional expectation and variance.
- It is defined for all $h > 0$ and not only for $h = 1$. Causality at different h does not imply each other. They are neither sufficient nor necessary.
- Ω_t contains all the information in the universe up to time t that excludes the potential ignored common factors problem. The question is: how to measure Ω_t in practice? The unobserved common factors are always a potential problem for any finite information set.
- Instantaneous causality $\Omega_{t+h} - x_{t+h}$ and feedback is difficult to interpret unless one has additional structural information.

A refined definition

X_t does not Granger cause Y_{t+h} relative to information J_t, if

$$E(Y_{t+h}|J_t, X_t) = E(Y_{t+h}|J_t)$$

Remark: Note that causality here is defined as *relative to*. In other words, no effort is made to find the complete causal path and possible common factors.

Equivalent definition

For a l-dimension stationary process, Z_t, there exists a canonical MA representation

$$\begin{aligned} Z_t &= \mu + \Phi(B)u_t \\ &= \mu + \sum_{i=1}^{\infty} \Phi_i u_{t-i}, \Phi_0 = I_l \end{aligned}$$

A necessary and sufficient condition for variable k not to Granger-cause variable j is that $\Phi_{jk,i} = 0, \quad \text{for } i = 1, 2, \cdots$. If the process is invertible, then

$$\begin{aligned} Z_t &= C + A(B)Z_{t-1} + u_t \\ &= C + \sum_{i=1}^{\infty} A_i Z_{t-i} + u_t \end{aligned}$$

For two variables, or two-group of variables, j and k, variable k not to Granger-cause variable j iff $A_{jk,i} = 0, \quad \text{for } i = 1, 2, \cdots$. The condition is good for all forecast horizon, h.

For a VAR(1) process with dimension equal or greater than 3, $A_{jk,i} = 0, \quad \text{for } i = 1, 2, \cdots$ is sufficient for non-causality at $h = 1$ but insufficient for $h > 1$.

$$\begin{bmatrix} y_{1t} \\ y_{2t} \\ y_{3t} \end{bmatrix} = \begin{bmatrix} .5 & 0 & 0 \\ .1 & .1 & .3 \\ 0 & .2 & .3 \end{bmatrix} \begin{bmatrix} y_{1t-1} \\ y_{2t-1} \\ y_{3t-1} \end{bmatrix} + \begin{bmatrix} u_{1t} \\ u_{2t} \\ u_{3t} \end{bmatrix}$$

Then,

$$y_0 = \begin{bmatrix} u_{10} \\ u_{20} \\ u_{30} \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}; y_1 = A_1 y_0 = \begin{bmatrix} .5 \\ .1 \\ 0 \end{bmatrix};$$

$$y_2 = A_1^2 y_0 = \begin{bmatrix} .25 \\ .06 \\ .02 \end{bmatrix}$$

Summary

1. For bivariate or two groups of variables, IR analysis is equivalent to applying Granger-causality test to VAR model;
2. For testing the impact of one variable on the other within a high dimensional (≥ 2) system, IR analysis can not be substituted by the Granger-causality test. For example, for an VAR(1) process with dimension greater than 2, it does not suffice to check the upper right-hand corner element of the coefficient matrix in order to determine if the last variable is noncausal for the first variable. Test has to be based upon IR.

See Lutkepohl(1993) and Dufor and Renault (1998) for detailed discussion.

Impulse response and causal ordering

Residuals from a VAR model are generally correlated and applying the Cholesky decomposition is equivalent to assuming recursive causal ordering from the top variable to the bottom variable.

Changing the order of the variables could greatly change the results of the impulse response analysis.

Causal analysis for bivariate VAR

For a bivariate system, y_t, x_t defined by

$$\begin{bmatrix} y_t \\ x_t \end{bmatrix} = \begin{bmatrix} A_{11}(B) & A_{12}(B) \\ A_{21}(B) & A_{22}(B) \end{bmatrix} \begin{bmatrix} y_{t-1} \\ x_{t-1} \end{bmatrix} + \begin{bmatrix} u_{yt} \\ u_{xt} \end{bmatrix}$$
$$= \begin{bmatrix} \Phi_{11}(B) & \Phi_{12}(B) \\ \Phi_{21}(B) & \Phi_{22}(B) \end{bmatrix} \begin{bmatrix} u_{yt-1} \\ u_{xt-1} \end{bmatrix} + \begin{bmatrix} u_{yt} \\ u_{xt} \end{bmatrix}$$

x_t does not Granger-cause y_t if $\Phi_{12}(B) = 0$ or
$\Phi_{12,i} = 0$, for $i = 1, 2, \cdots$.
This condition is equivalent to $A_{12,i} = 0$, for $i = 1, 2, \cdots$,p.
The restrictions that all cross-lags coefficients are all zeros which can be tested by Wald statistics.

Bivariate AR(1) process $A_{ij}(B) = A_{ij}, i, j = 1, 2$
Four possible causal directions between x and y are:

1. Feedback, $H_0, x \leftrightarrow y$

$$H_0 = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$$

2. Independent, $H_1 : x \perp y$

$$H_1 = \begin{pmatrix} A_{11} & 0 \\ 0 & A_{22} \end{pmatrix}$$

3. x causes y but y does not cause $x, H_2, y \not\mapsto x$

$$H_2 = \begin{pmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{pmatrix}$$

4. y causes x but x does not cause $y, H_3, x \not\mapsto y$

$$H_3 = \begin{pmatrix} A_{11} & 0 \\ A_{21} & A_{22} \end{pmatrix}$$

Caines, Keng and Sethi(1981) two-stage testing procedure
First stage: test H_1 (null) against H_0, H_2 (null) against H_0, and H_3 (null) against H_0.
Second stage (if necessary): test H_1 (null) against H_2, and H_1 (null) against H_3.

Causal analysis for multivariate VAR

Possible causal structure grows exponentially as number of variables increase. Pairwise causal structure might change when different conditioning variables are added.

Caines, Keng and Sethi (1981) procedure

1. For a pair (X, Y), construct bivariate VAR with order chosen to minimize multivariate final prediction error (MFPE);
2. Apply the stagewise procedure to determine the causal structure of X, Y.
3. If a process X, has n multiple causal variables, $y_1, \ldots, y_n$, rank these variables according to the decreasing order of their specific gravity which is the inverse of $MFPE(X, y_i)$.
4. For each caused variable process, X, first construct the optimal univariate AR model using FPE to determine the lag order. The, add the causal variable, one at a time according to their causal rank and use FPE to determine the optimal orders at each step. Finally, we get the optimal ordered univariate multivariable AR model of X against its causal variables.
5. Pool all the optimal univariate AR models above and apply the Full Information Maximum Likelihood (FIML) method to estimate the system. Finally perform the diagnostic checking with the whole system as maintained model.

Causal analysis for Vector ARMA model

Let X be $n \times 1$ vector generated by

$$\Phi(B)X_t = \Theta(B)a_t$$

X_i does not cause X_j if and only if

$$det(\Phi_i(z), \Theta_{(j)}(z)) = 0$$

where $\Phi_i(B)$ is the ith column of the matrix $\Phi(z)$ and $\Theta_{(j)}(z)$ is the matrix $\Theta(z)$ without its jth column.

Bivariate (two-group) case

$$\begin{pmatrix} \Phi_{11}(B) & \Phi_{12}(B) \\ \Phi_{21}(B) & \Phi_{22}(B) \end{pmatrix} \begin{pmatrix} X_{it} \\ X_{2t} \end{pmatrix} = \begin{pmatrix} \Theta_{11}(B) & \Theta_{12}(B) \\ \Theta_{21}(B) & \Theta_{22}(B) \end{pmatrix} \begin{pmatrix} a_{1t} \\ a_{2t} \end{pmatrix}$$

Then, X_i does not cause X_j if and only if

$$\Phi_{21}(z) - \Theta_{21}(z)\Theta_{11}^{-1}(z)\Phi_{11}(z) = 0$$

If $n_1 = n_2 = 1$, Then, X_i does not cause X_j if and only if

$$\Theta_{11}(z)\Phi_{12}(z) - \Theta_{21}(z)\Phi_{11}(z) = 0$$

General testing procedures

1. Build a multivariate ARMA model for X_t,
2. Derive the noncausalisty conditions in term of AR and MA parameters, say $R_j(\beta_I) = 0, j = 1, \ldots, K$
3. Choose a test criterion, Wald, LM or LR test.

Let

$$T(\hat{\beta}_I) = (\frac{\partial R_j(B)}{\partial \beta_I}|_{\beta_I - \hat{\beta}_I})_{k \times k}$$

Let $V(\beta_I)$ be the asymptotic covariance matrix of $\sqrt{N}(\hat{\beta}_I = \beta_I)$. Then the Wald and LR test statistics are:

$$\begin{aligned} \xi_W &= NR(\hat{\beta}_I)'[T(\hat{\beta}_I)'V(\hat{\beta}_I)T(\hat{\beta}_I)]^{-1}R(\hat{\beta}_I), \\ \xi_{LR} &= 2(L(\hat{\beta}, X) - L(\hat{\beta}*, X)) \end{aligned}$$

where $\hat{\beta}*$ is the MLE of β under the constraint of noncausality.

ARMA(1,1)case

To illustrate, let X_t be a invertible 2-dimensional ARMA(1,1) model.

$$\begin{pmatrix} 1-\phi_{11}B & -\phi_{12}B \\ -\phi_{21}B & 1-\phi_{22}B \end{pmatrix} \begin{pmatrix} X_{1t} \\ X_{2t} \end{pmatrix} = \begin{pmatrix} 1-\theta_{11}B & \theta_{12}B \\ \theta_{21}B & \theta_{22}B \end{pmatrix} \begin{pmatrix} a_{1t} \\ a_{2t} \end{pmatrix}$$

X_1 does not cause X_2 if and only if

$$\Theta_{11}(z)\Phi_{21}(z) - \Theta_{21}(z)\Phi_{11}(z) = 0$$
$$(\phi_{21} - \theta_{21})z + (\theta_{11}\theta_{21} - \phi_{21}\theta_{11})z^2 = 0$$
$$\phi_{21} - \theta_{21} = 0, \quad \phi_{11}\theta_{21} - \phi_{21}\theta_{11} = 0$$

For the vector, $\beta_I = (\phi_{11}, \phi_{21}, \theta_{11}, \theta_{21})'$, the matrix

$$T(\beta_I) = \begin{pmatrix} 0 & \theta_{21} \\ 1 & -\theta_{11} \\ 0 & -\phi_{21} \\ -1 & \phi_{11} \end{pmatrix}$$

might not be nonsingular under the null of $H_0 : X_1$ does not cause X_2.

Remarks:

- The conditions are weaker than $\phi_{21} = \theta_{21} = 0$
- $\phi_{21} - \theta_{21} = 0$ is a necessary condition for $H_0, \phi_{21} = \theta_{21} = 0$ is sufficient condition and $\phi_{21} - \theta_{21} = 0, \& \phi_{11} = \theta_{11}$ are sufficient for H_0.

Let $H_0 : X_1$ does not cause X_2. Consider the following hypotheses:

$$\begin{aligned} H_0^1 &: \phi_{21} - \theta_{21} = 0; \\ H_0^2 &: \phi_{21} = \theta_{21} = 0 \\ H_0^3 &: \phi_{21} \neq 0, \phi_{21} - \theta_{21} = 0, \text{ and } \phi_{11} - \theta_{11} = 0 \\ \tilde{H}_0^3 &: \phi_{11} - \theta_{11} = 0 \end{aligned}$$

Then, $H_0^3 = \tilde{H}_0^3 \cap H_0^1, H_0^2 \subseteq H_0 \subseteq H_0^1, H_0^3 \subseteq H_0 \subseteq H_0^1$.

Testing procedures

1. Test H_0^1 at level α_1. If H_0^1 is rejected, then H_0 is rejected. Stop.
2. If H_0^1 is not rejected, test H_0^2 at level α_2. If H_0^2 is not rejected, H_0 cannot be rejected. Stop
3. If H_0^2 is rejected, test $\tilde{H}_0^3 : \phi_{11} - \theta_{11} = 0$ at level α_2. If $\tilde{H}_0^3$ is rejected, then H_0 is also rejected. If $\tilde{H}_0^3$ is not reject ed, then H_0 is also not rejected.

Causal analysis for nonstationary processes

The asymptotic normal or χ^2 distribution in previous section is build upon the assumption that the underlying processes X_t is stationary. The existence of unit root and cointegration might make the traditional asymptotic inference invalid.

Unit root

What is unit root?

The time series y_t as defined in $A_p(B)y_t = C(B)\epsilon_t$ has an unit root if $A_p(1) = 0$, $C(1) \neq 0$.

Why do we care about unit root?

- For y_t, the existence of unit roots implies that a shock in ϵ_t has permanent impacts on y_t.
- If y_t has a unit root, then the traditional asymptotic normality results usually no longer apply. We need different asymptotic theorems.

What is cointegration?

When linear combination of two I(1) process become an I(0) process, then these two series are cointegrated.

Why do we care about cointegration?

- Cointegration implies existence of long-run equilibrium;
- Cointegration implies common stochastic trend;
- With cointegration, we can separate short- and long- run relationship among variables;
- Cointegration can be used to improve long-run forecast accuracy;
- Cointegration implies restrictions on the parameters and proper accounting of these restrictions could improve estimation efficiency.

Let Y_t be k-VAR(p) series with r cointegration vector $\beta(p \times r)$.

$$\begin{aligned}
A_p(B)Y_t &= U_t \\
\Delta Y_t &= \Pi Y_{t-1} + \sum_{i=1}^{p-1} \Gamma_i \Delta Y_{t-i} + \Phi D_t + U_t \\
Y_t &= C \sum_{i=1}^{t} (U_t + \Phi D_i) + C^*(B)(U_t + \Phi D_t) + P_{\beta_\perp} Y_0 \\
A_p(1) &= -\Pi = \alpha\beta' \\
C &= \beta_\perp(\alpha'_\perp \Gamma \beta_\perp)^{-1}\alpha'_\perp
\end{aligned}$$

- Cointegration introduces one additional causal channel (error correction term) for one variable to affect the other variables. Ignoring this additional channel will lead to invalid causal analysis.
- For cointegrated system, impulse response estimates from VAR model in level without explicitly considering cointegration will lead to incorrect confidence interval and inconsistent estimates of responses for long horizons.

Recommended procedures for testing cointegration:

1. Determine order of VAR(p). Suggest choosing the minimal p such that the residuals behave like vector white noise;
2. Determine type of deterministic terms: no intercept, intercept with constraint, intercept without constraint, time trend with constraint, time trend without constraint. Typically, model with intercept without constraint is preferred;
3. Use trace or λ_{max} tests to determine number of unit root;
4. Perform diagnostic checking of residuals;
5. Test for exclusion of variables in cointegration vector;
6. Test for weak erogeneity to determine if partial system is appropriate;
7. Test for stability;
8. Test for economic hypotheses that are converted to homogeneous restrictions on cointegration vectors and/or loading factors.

Unit root, Cointegration and causality

For a VAR system, X_t with possible unit root and cointegration, the usual causality test for the level variables could be misleading. Let $X_t = (X_{1t}, X_{2t}, X_{3t})'$ with n_1, n_2, n_3 dimension respectively. The VAR level model is:

$$\begin{aligned} X_t &= J(B)X_{t-1} + u_t \\ &= \sum_{i=1}^{k} J_i X_{t-i} + u_t \end{aligned}$$

The null hypothesis of X_3 does not cause X_1 cen be formulated as:

$$H_0 : J_{1,13} = J_{2,13} = \cdots = J_{k,13} = 0$$

Let F_{LS} be the Wald statistics for testing H_0.

1. If X_t has unit root and is not cointegrated, F_{LS} converges to a limiting distribution which is the sum of χ^2 and unit root distribution. The test is similar and critical values can be constructed. Yet, it is more efficient and easier to difference X_t and test causality for the differenced VAR.
2. If there is sufficient cointegration for X_3 then $F_{LS} \rightarrow \chi^2_{n_1 n_2 k}$. , More specifically, let $A = (A_1, A_2, A_3)$ be the cointegration vector. The usual asymptotic distribution results hold if $rank(A_3) = n_3$, ie. all X_3 appear in the cointegration vector.
3. If there is not sufficient cointegration, ie. not all X_3 appears in the cointegration vector, then the limiting distribution contain unit root and nuisance parameters.

Error correction model

$$\Delta X_t = J^*(B)\Delta X_{t-1} + \Gamma A' X_{t-1} + u_t$$

where Γ, A are respectively the loading matrix and cointegration vector. Partition Γ, A conforming to X_1, X_2, X_3.
Then, if $\text{rank}(A_3) = n_3$ or $\text{rank}(\Gamma_1) = n_1$, $F_{ML} \to \chi^2_{n_1 n_2 k}$.
Testing with ECM the usual asymptotic distribution hold when there are sufficient cointegrations or sufficient loading vector.
Remark: The Johansen test seems to assume sufficient cointergation or sufficient loading vector.

Toda and Yamamoto (1995)

Test of causality without pretesting cointegration.
For an VAR(p) process and each series is at most I(k), then estimate the augmented VAR(p+k) process even the last k coefficient matrix is zero.
$X_t = A_1 X_{t-1} + \cdots + A_p X_{t-k} + \cdots + A_{p+k} X_{t-p-k} + U_t$
and perform the usual Wald test $A_{kj,i} = 0, i = 1, \cdots, p$. The test statistics is asymptotical χ^2 with degree of freedom m being the number of constraints.
The result holds no matter whether X_t is I(0) or I(1) and whether there exist cointegration.

As there is no free lunch under the sun, the Toda-Yamamoto test suffer the following weakness.

- Inefficient as compared with ECM where cointegration is explicitly considered.
- Cannot distinguish between short run and long run causality.
- Cannot test for hypothesis on long run equilibrium, say PPP which is formulated on cointegration vector.

One more remark: Cointegration between two variables implies existence of long-run causality for at least one direction. Testing cointegration and causality should be considered jointly.

Causal analysis using graphical models

A directed graph assigns a contemporaneous causal flow among a set of variables based on correlations and partial correlations. The edge relationship of each pair of variables characterizes the causal relationship between them.

- No edge: (conditional) independence between two variables
- an undirected edge ($X - Y$): correlation with no particular causal interpretation.
- directed edge ($Y \rightarrow X$): Y causes X but X does not cause Y conditional upon other variables.
- bidirected edge ($X \leftrightarrow Y$): bidirectional causality between these two variables. There is contemporaneous feedback between X and Y.

Illustration

let X, Y, Z be three variables under investigation.
$Y \leftarrow X \rightarrow Z$ represents the fact that X is the common cause of Y and Z. Unconditional correlation between Y and Z is nonzero but conditional correlation between Y and Z given X is zero.
$Y \rightarrow X \leftarrow Z$ says that both Y and Z cause X. Thus, unconditional correlation between Y and Z is zero but conditional correlation between Y and Z given X is nonzero.
$Y \rightarrow X \rightarrow Z$ states the fact that Y causes X and X causes Z. Being conditional upon X, Y is uncorrelated with Z.

The direction of the arrow is then transformed into the zero constraints of $A(i,j), i \neq j$.
Let $u_t = (X_t, Y_t, Z_t)'$ and then the corresponding A matrix for the three cases discussed above denoted as A_1, A_2 and A_3 are:

$$A_1 = \begin{bmatrix} 1 & 0 & 0 \\ a_{21} & 1 & 0 \\ a_{31} & 0 & 1 \end{bmatrix}; A_2 = \begin{bmatrix} 1 & a_{12} & a_{13} \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}; A_3 = \begin{bmatrix} 1 & a_{12} & 0 \\ 0 & 1 & 0 \\ a_{31} & 0 & 1 \end{bmatrix}$$

Search algorithms

Several search algorithms are available and the PC algorithm seems to be the most popular one.

1. Start with a graph in which each variable is connected by an edge with every other variable.
2. Compute the unconditional correlation between each pair of variables and remove the edge for the insignificant pairs.
3. Compute the 1-*th* order conditional correlation between each pair of variables and eliminate the edge between the insignificant ones.
4. Repeat the procedure to compute the *i*-*th* order conditional correlation until $i = N$-2, where N is the number of variables under investigation.

Fisher's z statistic is used in the significance test:

$$z(i,j|K) = 1/2(n - |K| - 3)^{(1/2)} ln(\frac{|1 + r[i,j|K]|}{|1 - r[i,j|K]|})$$

where $r([i,j|K])$ denotes conditional correlation between variables, which i and j being conditional upon the K variables, and $|K|$ the

1. For each pair of variables (Y, Z) that are unconnected by a direct edge but are connected through an undirected edge through a third variable X, assign $Y \rightarrow X \leftarrow Z$ if and only if the conditional correlations of Y and Z conditional upon all possible variable combinations with the presence of the X variable are nonzero.
2. Repeat the above process until all possible cases are exhausted.
3. If $X \rightarrow Z$, $Z - Y$ and X and Y are not directly connected, assign $Z \rightarrow Y$.
4. If there is a directed path between X and Y (say $X \rightarrow Z \rightarrow Y$) and there is an undirected edge between X and Y, assign $X \rightarrow Y$.

Causality on spectral domain

Causality on time domain is qualitative but the strength of causality y at each frequency can be measured on spectral domain. An ideal model for analyzing permanent consumption theory.

Bivariate case

$$\begin{bmatrix} x_t \\ y_t \end{bmatrix} = \begin{bmatrix} \Lambda_{11}(B) & \Lambda_{12}(B) \\ \Lambda_{21}(B) & \Lambda_{22}(B) \end{bmatrix} \begin{bmatrix} e_{xt} \\ e_{yt} \end{bmatrix}$$

Rewrite the above as

$$\begin{bmatrix} x_t \\ y_t \end{bmatrix} = \begin{bmatrix} \Gamma_{11}(B) & \Gamma_{12}(B) \\ \Gamma_{21}(B) & \Gamma_{22}(B) \end{bmatrix} \begin{bmatrix} \tilde{e}_{xt} \\ \tilde{e}_{yt} \end{bmatrix}$$

where

$$\begin{bmatrix} \Gamma_{11}(B) & \Gamma_{12}(B) \\ \Gamma_{21}(B) & \Gamma_{22}(B) \end{bmatrix} = \begin{bmatrix} \Lambda_{11}(B) & \Lambda_{12}(B) \\ \Lambda_{21}(B) & \Lambda_{22}(B) \end{bmatrix} \begin{bmatrix} 1 & 0 \\ \rho & 1 \end{bmatrix}$$

and

$$\begin{bmatrix} \tilde{e}_{xt} \\ \tilde{e}_{yt} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ -\rho & 1 \end{bmatrix} \begin{bmatrix} e_{xt} \\ e_{yt} \end{bmatrix}$$

$$f_x(w) = \frac{1}{2\pi}\{|\Gamma_{11}(z)|^2 + |\Gamma_{12}(z)|^2(1-\rho^2)\}$$

where $z = e^{-iw}$.

Hosoya's measure of one-way causality

$$
\begin{aligned}
M_{y\rightarrow x}(w) &= log[\frac{f_x(w)}{1/2\pi|\Gamma_{11}(z)|^2}] \\
&= log[1+\frac{|\Lambda_{12}(z)|^2(1-\rho^2)}{|\Lambda_{11}(z)+\rho\Lambda_{12}(z)^2|}]
\end{aligned}
$$

Error correction model

Let x_t, y_t be I(1) and $u_t = y_t - Ax_t$ be an I(0). The the error correction model is:

$$\begin{aligned} \Delta x_t &= \lambda_1 u_{t-1} + a_1(B)\Delta x_{t-1} + b_1(B)\Delta y_{t-1} + e_{xt} \\ \Delta y_t &= \lambda_2 u_{t-1} + a_2(B)\Delta x_{t-1} + b_2(B)\Delta y_{t-1} + e_{yt} \end{aligned}$$

$$\begin{bmatrix} D(B)x_t \\ D(B)y_t \end{bmatrix} = \begin{bmatrix} (1-B)(1-b_2B)\lambda_2 B & \lambda_1 B + b_1 B(1-B) \\ (1-B)a_2B - \lambda_2 AB & \lambda_1 AB + (1-a_1B)(1-B) \end{bmatrix} \begin{bmatrix} e_{xt} \\ e_{yt} \end{bmatrix}$$

where $D(B)$ arises from matrix inversion. Then,

$$M_{y\to x}(w) = log[1 + \frac{|\lambda_1 + b_1(1-z)|^2(1-\rho^2)}{|\{(1-z)(\bar{z}-b_2) - \lambda_2\} + \{\lambda_1 + b_1(1-z)\}|\rho|^2}]$$

where $\bar{z} = e^{iw}$.

Don't Do

1. Don't do single equation causality testing and draw inference on the causal direction,
2. Don't test causality between each possible pair of variables and then draw conclusions on the causal directions among variables,
3. Do not employ the two-step causality testing procedure though it is not an uncommon practice.
   People often test for cointegration first and then treat the error-correction term as an independent regressor and then apply the usual causality testing. This procedure is flawed for the following reasons. First, EC term is estimated and using it as an regressor in the next step will give rise to generated regressor problem. That is, the usual standard deviation in the second step is not right. Second, there could be more than one cointegration vectors and linear combination of them are also cointegrated vectors.

Do

1. Examine the graphs first. Look for pattern, mismatch of seasonality, abnormality, outliers, etc.
2. Always perform diagnostic checking of residuals: Time series modelling does not obtain help from economic theory and depends heavily upon statistical aspects of correct model specification. Whiteness of residuals are the key assumption.
3. Often graph the residuals and check for abnormality and outliers.
4. Be aware of seasonality for data not seasonally adjusted.
5. Apply the Wald test within the Johansen framework where one can test for hypothesis on long- and short- run causality.
6. When you employ several time series methods or analyze several similar models, be careful about the consistency among them.
7. Always watch for balance between explained and explanatory variables in regression analysis. For example, if the dependent variable has a time-trend but explanatory variables are limited

Empirical examples

On-line demonstration of the following examples

1. Evaluating the effectiveness of interest rate policy in Taiwan: an impulse responses analysis
   Lin(2003a).
2. Modelling information flow among four stock markets in China
   Lin and Wu (2003).
3. Causality between export expansion and manufacturing growth
   Liang, Chou and Lin (1995).