Exploring Granger Causality for Time series via Wald Test on Estimated Models with Guaranteed Stability

Nuntanut Raksasri
Jitkomut Songsiri

Department of Electrical Engineering, Faculty of Engineering, Chulalongkorn University

Nov 3, 2016
OUTLINE

- Introduction and Objective
- Granger causality
- Statistical test for Granger causality Analysis
- Stability Conditions
- Result and Conclusion
By knowing the relationship of the parameters in time series data, we can explain the dynamic of the time series data.

- explain relationship between the time series data by using the Granger causality concept and autoregressive model.
- estimate stable model with Granger causality constraints.
Flow Chart of the Paper

1. **ML estimation of AR model with closed-form solution**
   
   - $\hat{A}, \Sigma$

2. **Wald test on $\hat{A}$**
   
   - Zero pattern of $\hat{A}$

3. **Estimation of AR Subject to Granger causality condition**
   
   - Sparsity pattern of $\hat{A}$

4. **Learn relationship**
   
   - $A$, $\Sigma$

5. **Estimation of AR Subject to Granger causality condition + Stability condition**
   
   - Sparsity pattern of $\hat{A}$

6. **Stable AR model with Granger causality**
   
   - All eigenvalues of $A$ lie on the unite circle

7. **AR model with Granger causality**
   
   - Some eigenvalues of $A$ lie outside the unit circle

8. **Is the model stable?**
   
   - no
   
   - yes

9. **Learn relationship**

10. **Model stability**
Time series data can be represented by an AR model,

$$y(t) = c + A_1 y(t - 1) + A_2 y(t - 2) + ... + A_p y(t - p) + v(t)$$  \hspace{1cm} (1)

- $y(t) = (y_1(t), y_2(t), ..., y_n(t)) \in \mathbb{R}^n$
- $A_1, A_2, ..., A_p \in \mathbb{R}^{n \times n}$ are AR coefficients ($p$ is lag order of the model)
- $c \in \mathbb{R}^n$ is a constant vector
- $v(t)$ is a Gaussian noise process with variance $\Sigma$
- the observations $y(1), y(2), ..., y(N)$ are available.
- $y(1), y(2), ..., y(p)$ are deterministic values and given.
“Granger causality” is a term for a specific notion of causality in time-series analysis. The idea of Granger causality is a simple one:

\[ X \xrightarrow{\text{G-causes}} Y \]

A variable \( X \) “Granger-causes” \( Y \) if \( Y \) can be better predicted using the histories of both \( X \) and \( Y \) than it can using the history of only \( Y \).
Apply the concept of Granger causality to AR model in equation (1) the causality of the model can be written in linear equation form that is if \( y_j \) “not Granger-cause” to \( y_i \) then

\[
(A_k)_{ij} = 0, \quad k = 1, 2, \ldots, p
\]

where \((A_k)_{ij}\) denotes the \((i, j)\) entry of \(A_k\).

Granger Causality structure can be read from the zero pattern of estimated AR coefficient matrix.
Consider AR(4) (AR model when p=4) and \( y(t) \in \mathbb{R}^3 \), if \( y_2 \) “not Granger-cause” to \( y_1 \) then the model could be

\[
\begin{bmatrix}
  y_1(t) \\
  y_2(t) \\
  y_3(t)
\end{bmatrix}
= c +
\begin{bmatrix}
  X & 0 & X \\
  X & X & X \\
  X & X & X
\end{bmatrix}
\begin{bmatrix}
  y_1(t-1) \\
  y_2(t-1) \\
  y_3(t-1)
\end{bmatrix}
+ 
\begin{bmatrix}
  X & 0 & X \\
  X & X & X \\
  X & X & X
\end{bmatrix}
\begin{bmatrix}
  y_1(t-2) \\
  y_2(t-2) \\
  y_3(t-2)
\end{bmatrix}
+ 
\begin{bmatrix}
  X & 0 & X \\
  X & X & X \\
  X & X & X
\end{bmatrix}
\begin{bmatrix}
  y_1(t-3) \\
  y_2(t-3) \\
  y_3(t-3)
\end{bmatrix}
+ 
\begin{bmatrix}
  X & 0 & X \\
  X & X & X \\
  X & X & X
\end{bmatrix}
\begin{bmatrix}
  y_1(t-4) \\
  y_2(t-4) \\
  y_3(t-4)
\end{bmatrix}
+ v(t)
\]
Maximum likelihood estimation

To estimate $A$ and $\Sigma$ by using maximum likelihood estimation, we solve the problem

$$\max_{A, \Sigma} \frac{N - p}{2} \log \det \Sigma^{-1} - \frac{1}{2} \|L(Y - AH)\|_F^2$$

(2)

when $L^T L = \Sigma^{-1}$ and the problem is equivalent to the least-squares problem

$$\min_{A} \|Y - AH\|_F^2$$

(3)

where

$$Y = \begin{bmatrix} y(p + 1) & y(p + 2) & \cdots & y(N) \end{bmatrix}_{n \times (N - p)} ,$$

$$H = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ y(p) & y(p + 1) & \cdots & y(N - 1) \\ y(p - 1) & y(p) & \cdots & y(N - 2) \\ \vdots & \vdots & \cdots & \vdots \\ y(1) & y(2) & \cdots & y(N - p) \end{bmatrix}_{(np+1) \times (N - p)}$$
Statistical test for Granger causality Analysis

The null hypothesis for Granger causality condition will be

\[ H_0 : (A_k)_{ij} = 0 \text{ for } k = 1, 2, \ldots, p \]

and the Wald test is based on the following test statistic:

\[ W_{ij} = \hat{B}_{ij}^T \left[ \widehat{\text{Avar}}(\hat{\theta})_{ij} \right]^{-1} \hat{B}_{ij} \]

where \( \hat{B}_{ij} = \left( (\hat{A}_1)_{ij}, (\hat{A}_2)_{ij}, \ldots, (\hat{A}_p)_{ij} \right) \), \( \hat{\theta} \) is the vectorization of \( \hat{A} \), and \( \widehat{\text{Avar}}(\hat{\theta})_{ij} \) is the main diagonal block of a consistent estimate of the asymptotic covariance matrix of \( \hat{\theta} \).
Statistical test for Granger causality Analysis

**IDEA**: if $H_0$ is true, $(\hat{A}_k)_{ij}$ should equal to 0

In Wald test, $H_0$ is reject if $W > C$

where $C = F^{-1}(1 - \alpha)$ is critical value

and $\alpha = \text{Prob}(W > C)$ is the significance level.

If $W_{ij} > C$

$C = 9.4877$ when $\alpha = 0.05$ and $p = 3$

Under the null hypothesis that $(A_k)_{ij} = 0$, the Wald statistic $W$ converges in distribution to Chi-square distribution with $p$ degrees of freedom.
Wald test Results

By generating \( y(t) = A H(t) + \nu(t) \) and the model parameter \( A_k \) are generated by choosing some element to be equal to zero.

\[
\begin{align*}
(a) \quad & \alpha = 0.01 \\
(b) \quad & \alpha = 0.1
\end{align*}
\]

□ is intersect between non-zeros components
〇 is true component is zero but the estimated is non-zero
＋ is true component is non-zero but the estimated is zero and blank is intersect between zero components
After knowing the Granger causality pattern, we solve this optimization problem to estimate $A$ with Granger causality condition.

$$\minimize_A \| Y - AH \|_F^2$$

subject to $$(A_k)_{ij} = 0$$

- The estimated model parameter is not guarantee to be stable
- we need a stability condition.
we can write the AR model in discrete-time linear system

\[
\begin{bmatrix}
  y(t) \\
  y(t-1) \\
  \vdots \\
  y(t-p+1)
\end{bmatrix}
= \begin{bmatrix}
  A_1 & A_2 & \cdots & A_{p-1} & A_p \\
  I & 0 & \cdots & 0 & 0 \\
  0 & I & \cdots & 0 & 0 \\
  \vdots & \vdots & \ddots & \vdots & \vdots \\
  0 & 0 & \cdots & I & 0 \\
\end{bmatrix}
\begin{bmatrix}
  y(t-1) \\
  y(t-2) \\
  \vdots \\
  y(t-p+1) \\
\end{bmatrix}
\]

The system is stable if and only if \( \max_i |\lambda_i(A)| < 1 \). The characteristic polynomial have \( A_k \) as a coefficient so the condition will be nonlinear in \( A \).
## Previous Works of Stability condition

<table>
<thead>
<tr>
<th>Researcher</th>
<th>Method</th>
<th>Difference in paper</th>
</tr>
</thead>
<tbody>
<tr>
<td>J. Mari et al., 2000</td>
<td>Apply Lyapunov theory to vector ARMA</td>
<td>Too complicated</td>
</tr>
<tr>
<td>S. L. Lacy et al., 2003</td>
<td></td>
<td></td>
</tr>
<tr>
<td>V. Cerone et al., 2010</td>
<td>Jury’s test to guarantee BIBO stability on SISO LTI system</td>
<td>MIMO linear system</td>
</tr>
<tr>
<td>L. Buesing et al., 2012</td>
<td>Apply Lyapunov theory on LDS model to estimate $A$ that comply with $A^T A \prec I$</td>
<td>structure of $A$</td>
</tr>
<tr>
<td>K. Turksoy et al., 2013</td>
<td>Use Gershgorin circle theory on ARMAX</td>
<td>similar</td>
</tr>
</tbody>
</table>
Sufficient Condition for stability

Spectral radius and Induced norm

From spectral radius $\rho(A) = \max_i |\lambda_i(A)|$ then the system is stable if $\rho(A) < 1$ and by the inequality

$$\rho(A) \leq \|A\|$$

if assume that $\|A\| < 1$ it will affect $\rho(A) < 1$ when $\|A\|$ is a induced norm.

We choose the infinity-norm of $A$ to be the sufficient condition for stability.

$$\|A\|_{\infty} \leq 1$$

Due to structure of $A$

- $\|A\|_1 \leq 1$ and $\|A\|_2 \leq 1$ lead to meaningless which is $A_1, \ldots, A_{p-1}$ are equal to zero.
- $\|A\|_F \leq 1$ is impossible
We can estimate a stable model parameter with Granger causality condition by solving this problem.

\[
\min_{\mathbf{A}} \quad \| \mathbf{Y} - \mathbf{A} \mathbf{H} \|_F^2 \\
\text{subject to} \quad \mathbf{A} = \begin{bmatrix}
\mathbf{A}_1 & \mathbf{A}_2 & \ldots & \mathbf{A}_{p-1} & \mathbf{A}_p \\
\mathbf{I} & 0 & \ldots & 0 & 0 \\
0 & \mathbf{I} & \ldots & 0 & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & \ldots & \mathbf{I} & 0
\end{bmatrix},
\|\mathbf{A}\|_\infty \leq 1,
(A_k)_{ij} = 0, \quad (i,j) \in \{1, \ldots, n\} \times \{1, \ldots, n\}
\]

The problem is convex in quadratic form and can be solved by many solver.
Figure 1: Positions of the eigenvalues of \( A \) on complex plane

(c) no constraint

(d) with Granger causality constraints
Figure 2: Positions of the eigenvalues of $\mathcal{A}$ on complex plane with Granger causality and stability constraints
Conclusion

Time Series Data

ML estimation of AR model with closed-form solution

Estimation of AR Subject to Granger causality condition + Stability condition

ML estimation of AR model with closed-form solution

Wald test on \( \hat{A} \)

Zero pattern of \( \hat{A} \)

Stable AR model with Granger causality

Estimation of AR Subject to Granger causality condition

Sparsity pattern of \( \hat{A} \)

All eigenvalues of \( \mathcal{A} \) lie on the unite circle

Sparsity pattern of \( \hat{A} \)

Some eigenvalues of \( \mathcal{A} \) lie outside the unit circle

AR model with Granger causality

Is the model stable?

yes

no
Why we have to check the stability?

\[ \mathcal{A} \mid |\lambda_{max}(\mathcal{A})| < 1 \]

\[ \mathcal{A} \mid \|\mathcal{A}\|_\infty < 1 \]

Estimated \( \mathcal{A} \) without stability condition

Estimated \( \mathcal{A} \) with stability condition
L. Buesing, J. H. Macke, and M. Sahani.
Learning stable, regularised latent models of neural population dynamics.

V. Cerone, D. Piga, and D. Rehruto.
Bounding the parameters of linear systems with stability constraints.

W. H. Greene.
Econometric analysis, 2008.

J. Mari, P. Stoica, and T. McKelvey.
Vector arma estimation: A reliable subspace approach.
S. L. Lacy and D. S. Bernstein.  
Subspace identification with guaranteed stability using constrained optimization.  

J. Songsiri.  
Sparse autoregressive model estimation for learning granger causality in time series.  

K. Turksoy, E. S. Bayrak, L. Quinn, E. Littlejohn, and A. Cinar.  
Guaranteed stability of recursive multi-input-single-output time series models.  
Q&A
THANK YOU