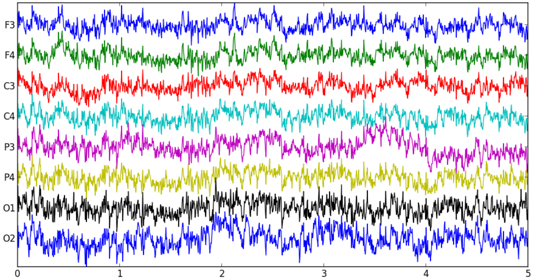
Exploring Granger causality for time series via statistical test on estimated models with guarantee stability

Nuntanut Raksasri
Adviser: Jitkomut Songsiri

Department of Electrical Engineering, Faculty of Engineering,
Chulalongkorn University

May 17, 2016

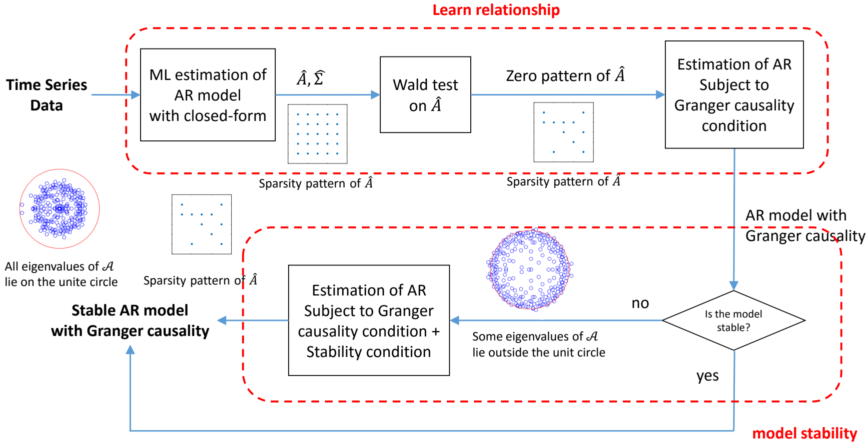
Introduction



By knowing the relationship of the parameters in time series data, we can explain the dynamic of the time series data.

In this project, we aim to explain relationship between the time series data by using the Granger causality concept and autoregressive model and estimate stable model with Granger causality constraints.

Flow Chart of the Project



Auto-Regressive (AR) Model

Time series data can be represented by an AR model,

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

- $y(t) = (y_1(t), y_2(t), ..., y_n(t)) \in \mathbb{R}^n$
- $A_1, A_2, ..., A_p \in \mathbb{R}^{n \times n}$ 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 Σ

Suppose the observations $y(1), y(2), ..., y(N)$ are available. We will assume that $y(1), y(2), ..., y(p)$ are deterministic values and given.

Granger causality

"Granger causality" is a term for a specific notion of causality in time-series analysis. The idea of Granger causality is a simple one:

$$X \xrightarrow{G-causes} Y$$

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.

Granger Causality

When 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

$$(A_k)_{ij} = 0, \quad k = 1, 2, .., p$$

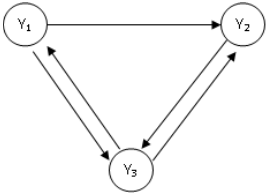
where $(A_k)_{ij}$ denotes the (i, j) entry of A_k.

- Granger Causality structure can be read from the zero pattern of estimated AR coefficient matrix.

example

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 Σ by using maximum likelihood estimation, we solving the problem

$$\underset{A,\Sigma}{\text{maximize}} \quad \frac{N-p}{2} \log\det \Sigma^{-1} - \frac{1}{2} \|L(Y - AH)\|_F^2 \tag{2}$$

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

$$\underset{A}{\text{minimize}} \quad \|Y - AH\|_F^2 \tag{3}$$

for estimating an A, when Y, H are the matrix that contain data of $y(t)$. The closed-form solution of A and Σ are

$$\hat{A} = YH^T(HH^T)^{-1}$$

$$\hat{\Sigma} = \frac{1}{N-p} \sum_{t=p+1}^{N} (y(t) - \hat{A}H(t))(y(t) - \hat{A}H(t))^T$$

Statistical test for Granger causality Analysis

The null hypothesis for Granger causality condition will be

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

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

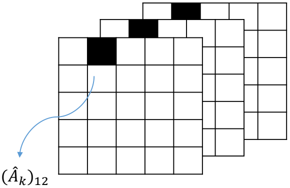
$$W_{ij} = \hat{B}_{ij}^T \left[\widehat{\textbf{Avar}}(\hat{\theta})_{ij} \right]^{-1} \hat{B}_{ij}$$

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 $\widehat{\textbf{Avar}}(\hat{\theta})_{ij}$ 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 α = **Prob**$(W > C)$ is the significance level.



If $W_{ij} >$ C → Non Zero

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

W

X	3.2486	35.0241	0.1986	11.5292
0.1279	X	153.831	35.5729	3.2287
13.4569	0.608	X	15.0093	2.1324
12.3471	21.6202	30.7289	X	2.2625
43.4160	2.3035	2.2536	1.5259	X

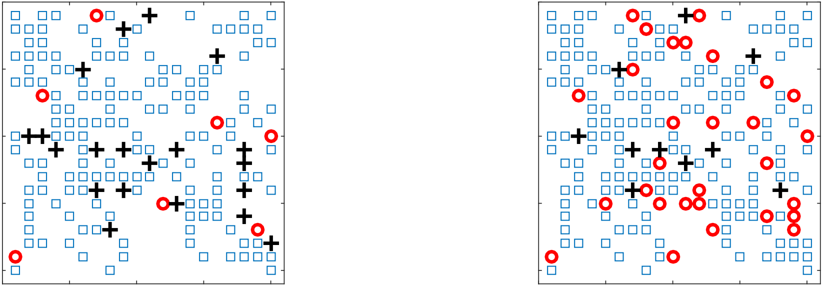
Zeros Structure

X	0	X	0	X
0	X	X	X	0
X	0	X	X	0
X	X	X	X	0
X	0	0	0	X

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 $y(t) = AH(t) + v(t)$ and the model parameter A_k are generated by choosing some element to be equal to zero.



(a) $\alpha = 0.01$ (b) $\alpha = 0.1$

□ is intersect between non-zeros components
◯ 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

AR estimation with Granger causality constraints

After we know the Granger causality pattern, we solve this optimization problem

$$\begin{array}{ll} \underset{A}{\text{minimize}} & \|Y - AH\|_F^2 \\ \text{subject to} & (A_k)_{ij} = 0 \end{array}$$

and we can estimate A with Granger causality condition.

However, the estimated model parameter is not guarantee to be stable so we need the stability condition.

Stability Condition

we can write the AR model in discrete-time linear system

$$\begin{bmatrix} y(t) \\ y(t-1) \\ \vdots \\ y(t-p+1) \end{bmatrix} = \underbrace{\begin{bmatrix} A_1 & A_2 & \dots & A_{p-1} & A_p \\ I & 0 & \dots & 0 & 0 \\ 0 & I & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & I & 0 \end{bmatrix}}_{\mathcal{A}} \begin{bmatrix} y(t-1) \\ y(t-2) \\ \vdots \\ y(t-p+1) \\ y(t-p) \end{bmatrix}$$

The system is stable if and only if $\max_i |\lambda_i(\mathcal{A})| < 1$. The characteristic polynomial have A_k as a coefficient so the condition will be nonlinear in A

Sufficient Condition for stability

Spectral radius and Induced norm

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

$$\rho(\mathcal{A}) \leq \|\mathcal{A}\|$$

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

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

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

Due to structure of $\mathcal{A}$

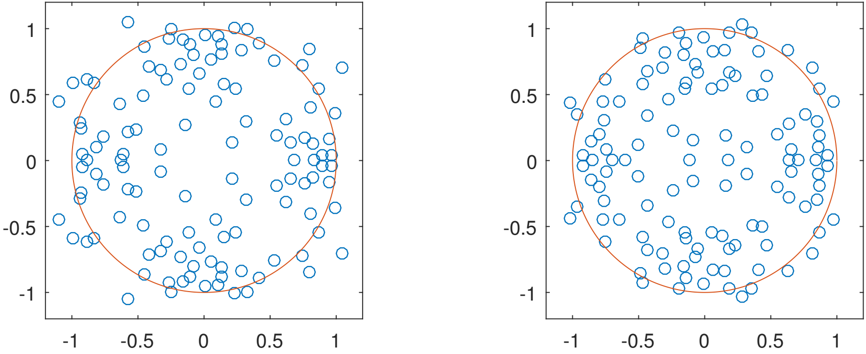
- $\|\mathcal{A}\|_1 \leq 1$ and $\|\mathcal{A}\|_2 \leq 1$ lead to meaningless
- $\|\mathcal{A}\|_F \leq 1$ is impossible

We can estimate a stable model parameter with Granger causality condition by solving this problem.

$$
\begin{array}{ll}
\underset{A}{\text{minimize}} & \|Y - AH\|_F^2 \\
\text{subject to} & \mathcal{A} = \begin{bmatrix} A_1 & A_2 & \dots & A_{p-1} & A_p \\ I & 0 & \dots & 0 & 0 \\ 0 & I & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & I & 0 \end{bmatrix}, \\
& \|\mathcal{A}\|_\infty \leq 1, \\
& (A_k)_{ij} = 0, \qquad (i,j) \subset \{1,\dots,n\}x\{1,\dots,n\}
\end{array}
$$

The problem is convex in quadratic form and can be solve by many solver.

Results



(c) no constraint

(d) with Granger causality constraints

Figure 1: Positions of the eigenvalues of $\mathcal{A}$ on complex plane

Results

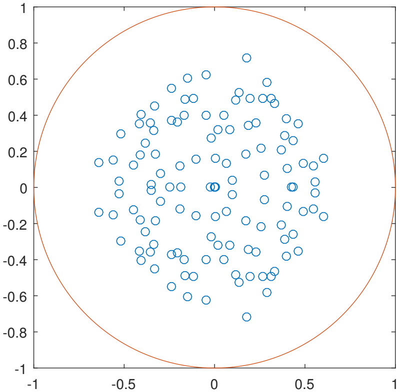
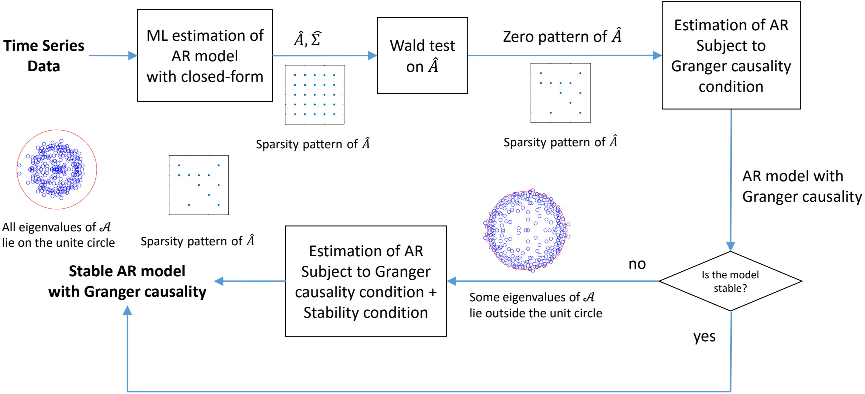
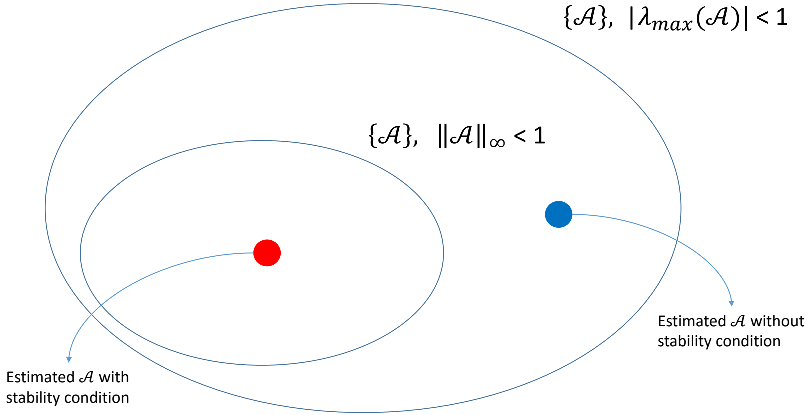


Figure 2: Positions of the eigenvalues of $\mathcal{A}$ on complex plane with Granger causality and stability constraints

Conclusion


Time Series Data
ML estimation of AR model with closed-form
$\hat{A}, \hat{\Sigma}$
Sparsity pattern of $\hat{A}$
Wald test on $\hat{A}$
Zero pattern of $\hat{A}$
Sparsity pattern of $\hat{A}$
Estimation of AR Subject to Granger causality condition
AR model with Granger causality
Is the model stable?
yes
no
Some eigenvalues of $\mathcal{A}$ lie outside the unit circle
Estimation of AR Subject to Granger causality condition + Stability condition
Sparsity pattern of $\hat{A}$
All eigenvalues of $\mathcal{A}$ lie on the unite circle
Stable AR model with Granger causality

Why we have to check the stability?



reference

L. Buesing, J. H. Macke, and M. Sahani.
Learning stable, regularised latent models of neural population dynamics.
Network: Computation in Neural Systems, 23(1-2):24–47, 2012.

V. Cerone, D. Piga, and D. Regruto.
Bounding the parameters of linear systems with stability constraints.
In *American Control Conference (ACC), 2010*, pages 2152–2157. IEEE, 2010.

W. H. Greene.
Econometric analysis, 2008.
Upper Saddle River, NJ: Prentice Hall, 24:25–26.

J. Mari, P. Stoica, and T. McKelvey.
Vector arma estimation: A reliable subspace approach.
Signal Processing, IEEE Transactions on, 48(7):2092–2104, 2000.

S. L. Lacy and D. S. Bernstein.
Subspace identification with guaranteed stability using constrained optimization.
Automatic Control, IEEE Transactions on, 48(7):1259–1263, 2003.

J. Songsiri.
Sparse autoregressive model estimation for learning granger causality in time series.
In *Acoustics, Speech and Signal Processing (ICASSP), 2013 IEEE International Conference on*, pages 3198–3202. IEEE, 2013.

K. Turksoy, E. S. Bayrak, L. Quinn, E. Littlejohn, and A. Cinar.
Guaranteed stability of recursive multi-input-single-output time series models.
In *American Control Conference (ACC), 2013*, pages 77–82. IEEE, 2013.

Q&A

THANK YOU