Senior Project Proposal 2102490 Year 2019

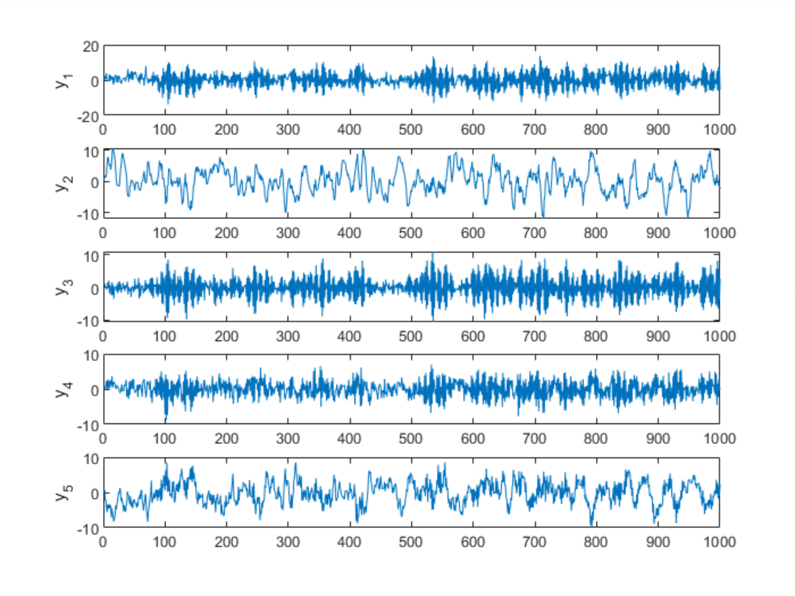
Learning Granger causality for multivariate time series using state-space models

Anawat Nartkulpat ID 5930562521

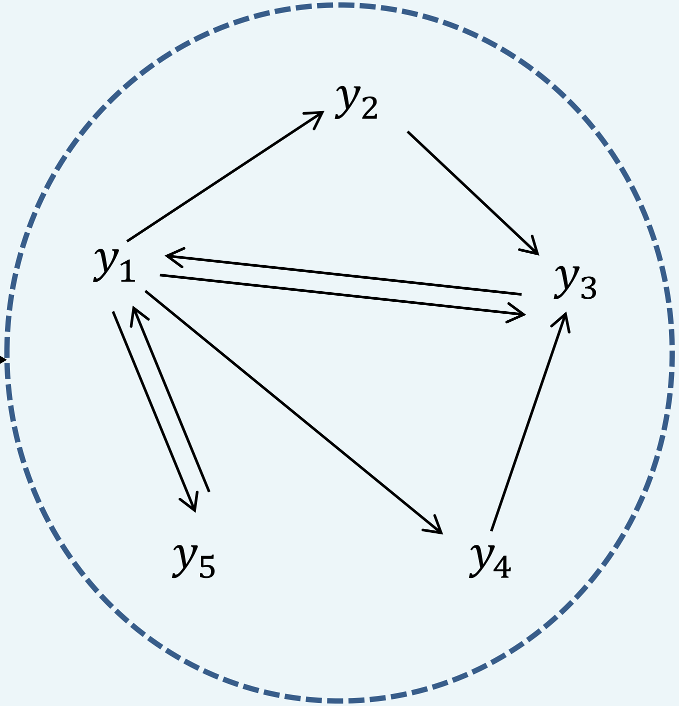
Advisor: Assist. Prof. Jitkomut Songsiri

Department of Electrical Engineering, Faculty of Engineering
Chulalongkorn University

Introduction

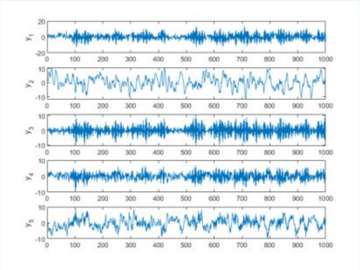


Time series data



Causality pattern

Introduction



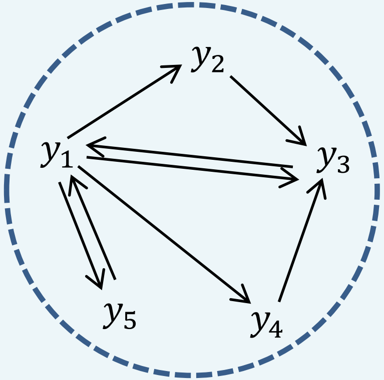
Model estimation

- Non-linear models
- Linear models
  - VAR
  - VARMA
  - State-space

Computing Granger causality

Significance test/classification of zero causality

- Clustering methods
  - GMM [1]
- Statistical methods
  - Parametric
  - Non-parametric
    - Permutation test
    - Bootstraps



Objectives

- To develop a scheme for classifying the zero patterns of the Granger causality of state-space models using the permutation test.

- To compare the performance of the permutation test with the Gaussian mixture models method in classification of zero and non-zero entries of Granger causality matrix obtained from state-space model

Methodology

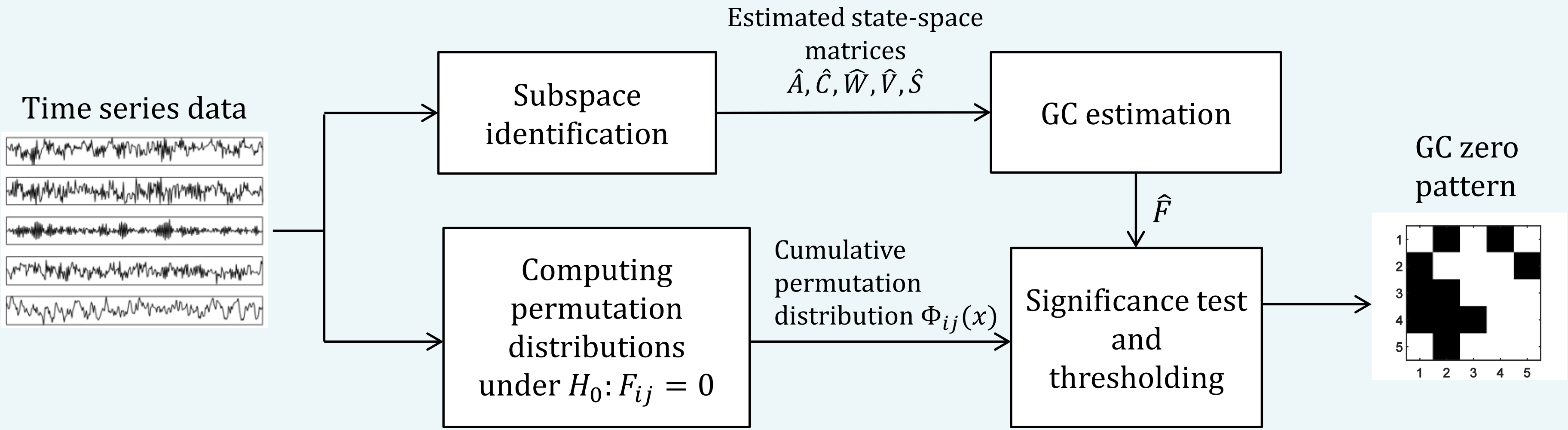


Figure 1. The scheme for learning GC pattern using state-space models.

Ground truth model

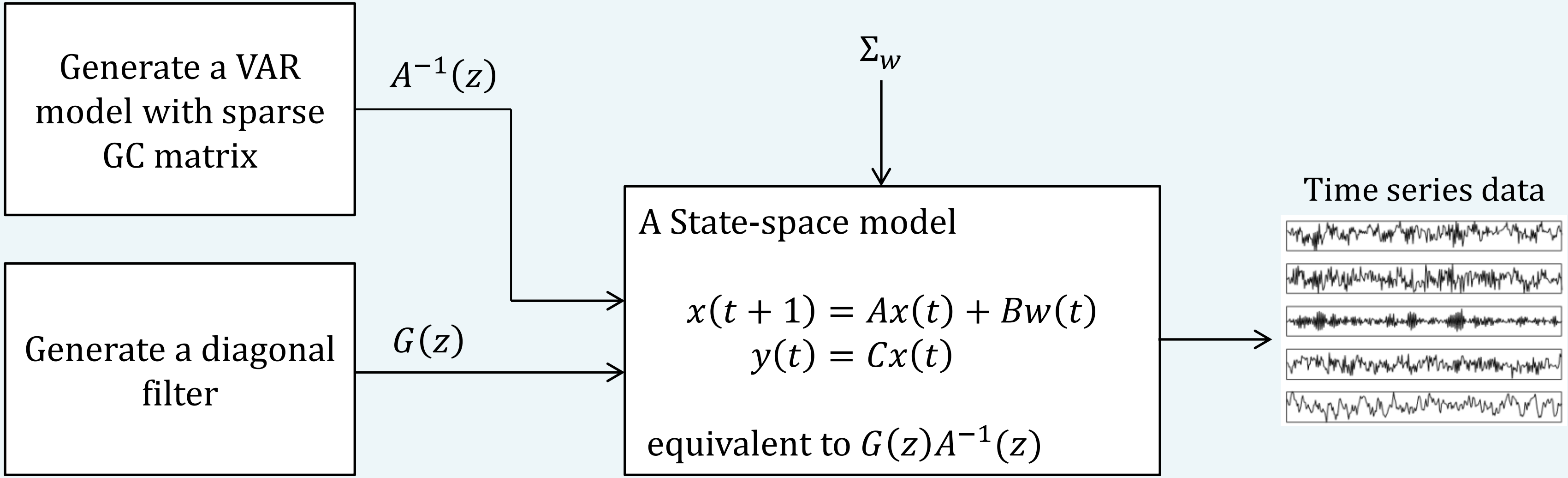


Figure 2. The scheme for generating ground truth model and time series data.

Subspace identification

We consider estimating parameters A, C, W, V, S of a stochastic state-space model

$$x(t+1) = Ax(t) + w(t)$$
$$y(t) = Cx(t) + v(t)$$

This method is based on orthogonal projection. Suppose that the outputs Y is known, it was shown in [2] that

$\mathcal{O}_i \equiv Y_{i|2i-1} / Y_{0|i-1} = Y_f / Y_p$ (Projecting the future outputs onto the past output space)

$\mathcal{O}_i = \Gamma_i \hat{X}_i \Longrightarrow \hat{X}_i = \Gamma_i^{\dagger} \mathcal{O}_i$ and $\hat{X}_{i+1} = \Gamma_{i-1}^{\dagger} \mathcal{O}_{i-1}$

$$\begin{bmatrix} \hat{X}_{i+1} \\ Y_{i|i} \end{bmatrix} = \begin{bmatrix} A \\ C \end{bmatrix} \hat{X}_i + \begin{bmatrix} \rho_w \\ \rho_v \end{bmatrix} \Longrightarrow \begin{cases} \begin{bmatrix} \hat{A} \\ \hat{C} \end{bmatrix} = \begin{bmatrix} \hat{X}_{i+1} \\ Y_{i|i} \end{bmatrix} \hat{X}_i^{\dagger} \\ \begin{bmatrix} \widehat{W} & \hat{S} \\ \hat{S}^T & \hat{V} \end{bmatrix} = \frac{1}{j} \begin{bmatrix} \rho_w \\ \rho_v \end{bmatrix} \begin{bmatrix} \rho_w \\ \rho_v \end{bmatrix}^T \end{cases}$$

Granger causality of state-space model

The measure of the Granger causality from y_j to y_i is defined by

$$F_{ij} = \log \frac{\Sigma_{ii}^R}{\Sigma_{ii}}$$

where Σ_{ii} is the covariance of the prediction error given all other y_k and Σ_{ii}^R is the covariance of the prediction error given all other y_k except y_j [3].

Together with this definition, we can define a Granger causality matrix

$$F = \begin{bmatrix} F_{11} & F_{12} & \cdots & F_{1n} \\ F_{21} & F_{22} & \cdots & F_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ F_{n1} & F_{n2} & \cdots & F_{nn} \end{bmatrix}$$

Granger causality of state-space model

For a state-space model

$$\begin{aligned} x(t+1) &= Ax(t) + w(t) \\ y(t) &= Cx(t) + v(t) \end{aligned}$$

where $\boldsymbol{E}\begin{bmatrix} w \\ v \end{bmatrix}\begin{bmatrix} w \\ v \end{bmatrix}^{\boldsymbol{T}} = \begin{bmatrix} W & S \\ S^T & V \end{bmatrix}$, we consider the reduced model with

$$\begin{aligned} x(t+1) &= Ax(t) + w(t) \\ y^R(t) &= C^R x(t) + v^R(t) \end{aligned}$$

where y^R is y without y_j and C^R is C without j^{th} row.

We have $\Sigma = \mathrm{CPC^T} + \mathrm{V}$ and $\Sigma_R = C^R P^R (C^R)^T + V^R$ where V^R is V without j^{th} row and column and P is solved from DARE [1]

$$P = APA^T - (APC^T + S)(CPC^T + V)^{-1}(CPA^T + S^T) + W$$

Permutation test

- The distribution of F_{ij} is unknown
- The null hypothesis H_0: $F_{ij} = 0$ is to be tested

Justification: Under the true null hypothesis, y_i is not Granger-caused by y_j, so rearranging data in channel y_j does not change the outcome.

So, we may form a distribution of F_{ij} under H_0 empirically from the permutations of data [4].

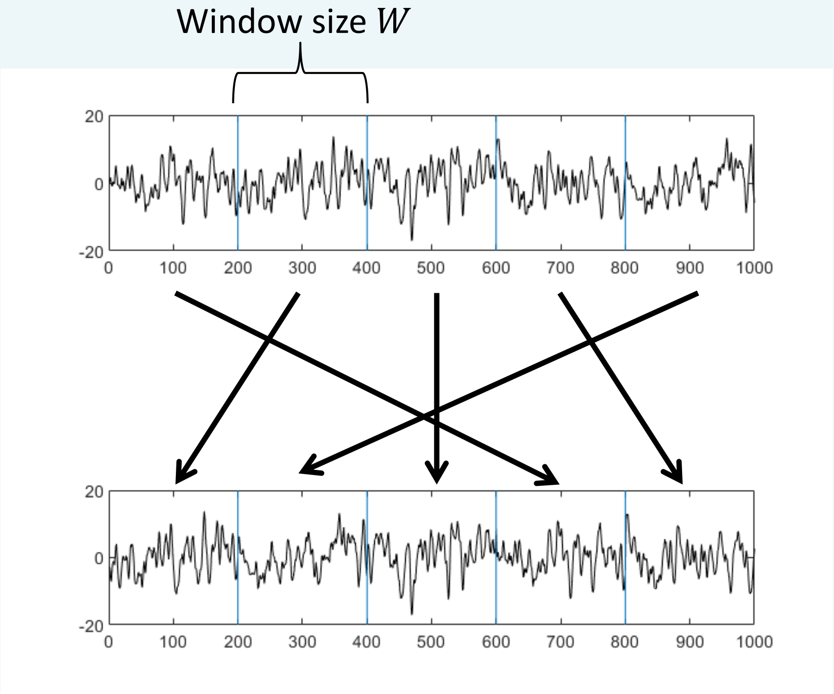


Figure 3. One of the possible permutations.

Permutation test

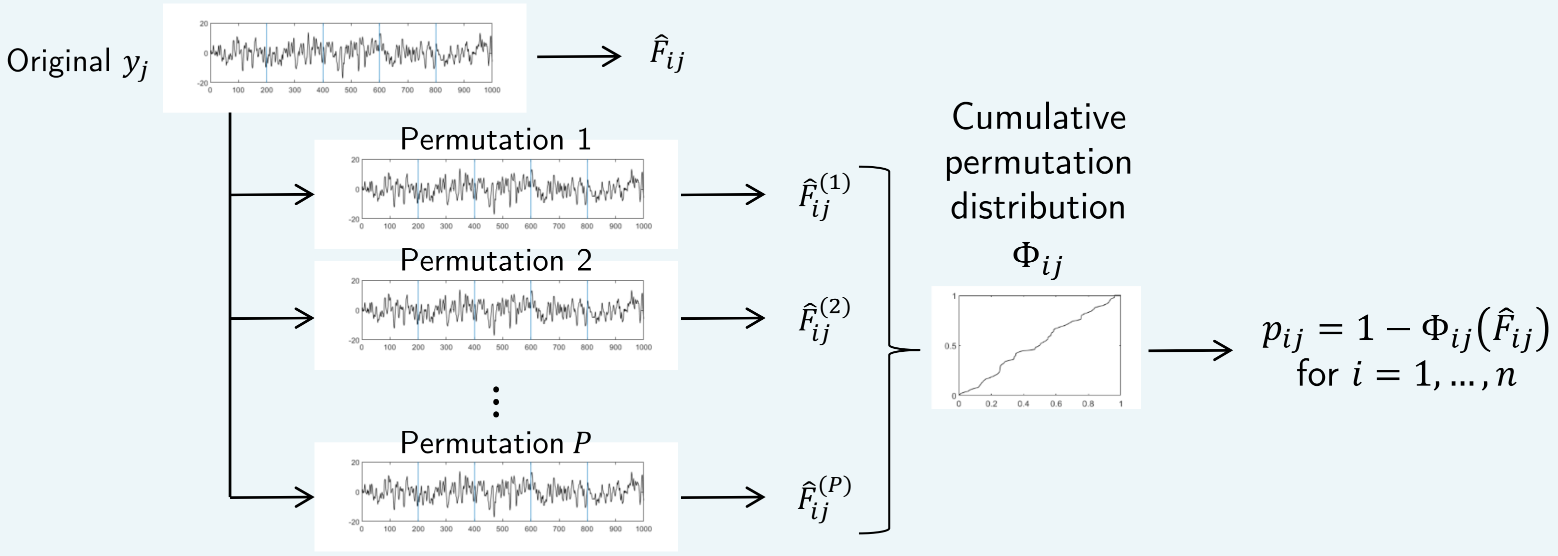


Figure 4. The scheme for calculating p-values.

Permutation test

The number of permutations (P) is the factorial of the number of windows. Since this number can be very large, we performed a Monte-Carlo permutation test in which the number of permutations used is smaller than the number of all possible permutations and the samples of rearrangement are drawn randomly [5].

Repeat for all j and the p-value matrix is then obtained

$$\begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1n} \\ p_{21} & p_{22} & \cdots & p_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ p_{n1} & p_{n2} & \cdots & p_{nn} \end{bmatrix}$$

For a given significance level α, F_{ij} can be tested to decide that $F_{ij} = 0$ or not by thresholding the p-values.

Preliminary results

An experiment was performed to study how the number of permutation (P) affects the performance of the permutation test

Hypothesis: The performance of the permutation test increase with the number of permutation

Control variable:
We generated 15 ground truth models with the following **specification**

- 5 channels
- 20 states (generate a VAR models of order 2 and a filter with 2 poles)
- 1000 time points for time series data

The estimation of state-space models were done with exactly 20 states.
Data were partitioned into 10 windows in permutation tests.

Preliminary results

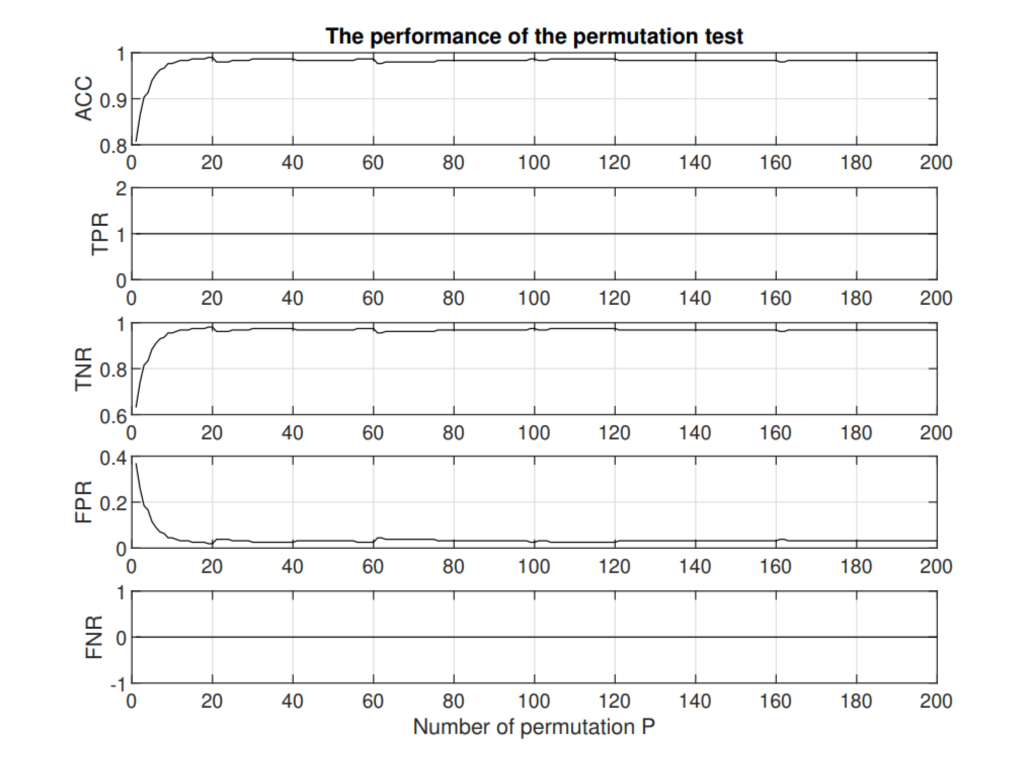


Figure 5. The performance of the permutation test.

Preliminary results

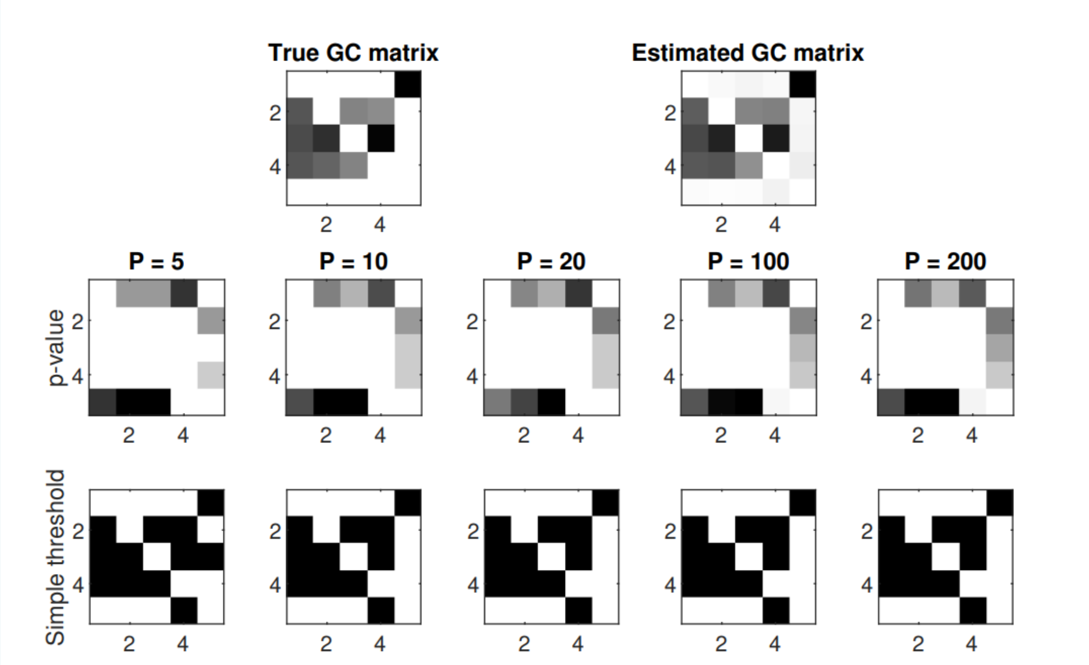


Figure 6. An example of the results after thresholding.

Project Planning

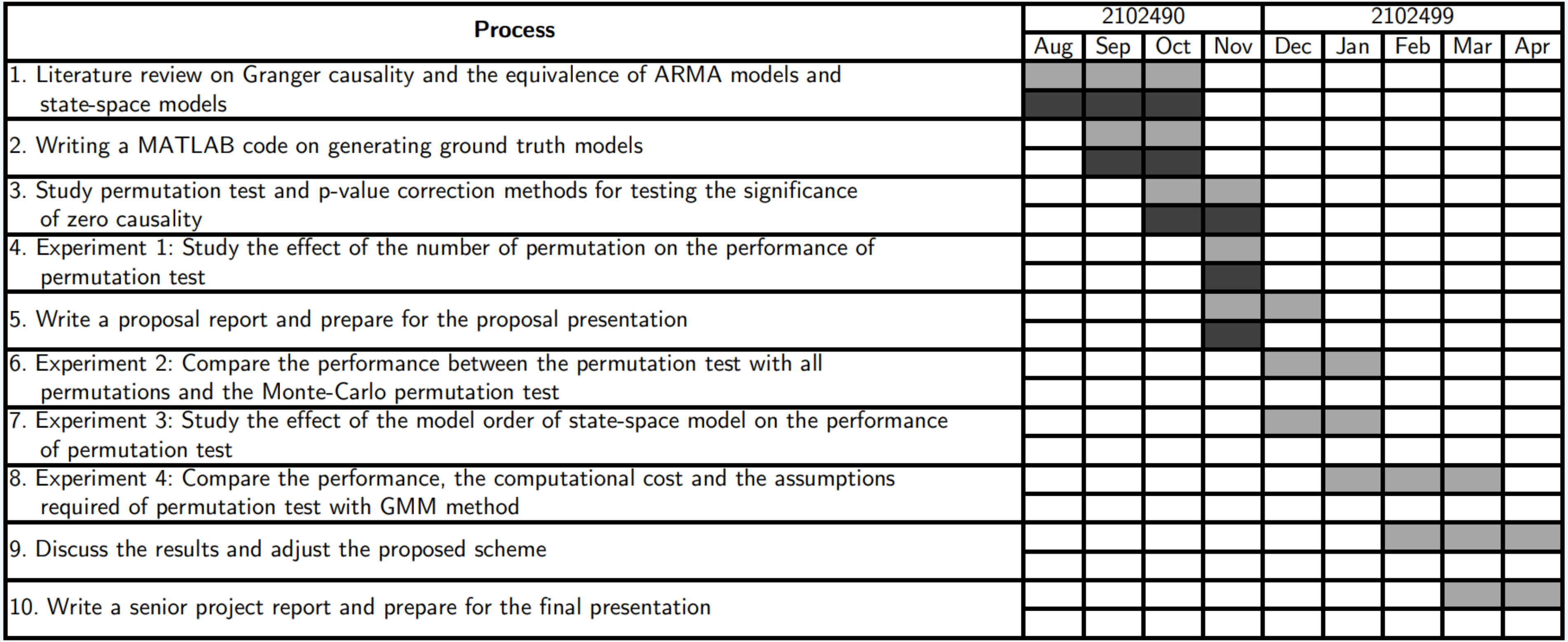


Figure 7. The Gantt chart of this project.

Project Planning

We plan to do the following experiments in the next semester.

- **Experiment 2**: Compare the performance of the permutation test with every permutations with the Monte-Carlo permutation.

  Experiment settings: With the same data, we use large window size ($W = 5$) for the normal permutation test so that the number of permutation is $5! = 120$ which is feasible. Then we compare the result with the Monte-Carlo permutation test.

- **Experiment 3**: Study the effect of the number of states used in subspace identification on the performance of the permutation test.

  Experiment settings: With the same data, we estimate state-space models with different assumption on the number of states and compare them.

Project Planning

- **Experiment 4**: Compare the performance and the computational cost of the permutation test with the GMM method

Experiment settings: With the same data, we compare the performance measures of the permutation test with the GMM method proposed in [1]. The parameters of the permutation test are chosen based on the previous result. The computation time is also compared.

Reference

[1] Songsiri, J., Learning brain connectivity from EEG time series, 2019

[2] Van Overschee P, De Moor BL. Subspace identification for linear systems: Theory—Implementation—Applications. Springer Science & Business Media; 2012.

[3] L. Barnett and A. K. Seth, “Granger causality for state space models,” Physical Review E, vol. 91, no. 4, 2015

[4] Thomas E Nichols and Andrew P Holmes. Nonparametric permutation tests for functional neuroimaging: a primer with examples. *Human brain mapping*, 15(1):1-25, 2002.

[5] Good, P., Permutation, parametric, and bootstrap tests of hypotheses. 2005.

Q&A