

Causal Relationship Approaches Comparison: Bayesian Network Inference and Granger Causality

Abstract

To explore the network structure of genes, proteins and neurons etc is one of the most fundamental biological process and data analysis in the system biology. In biological experiments, the recorded multi-variable time series (gene data, protein data and neuron data) are used to derive the relationship structure for all the measured genes, proteins or other variables. Currently, there are two main approaches which are often used to infer the causal relationships or interactions among a set of variables. One is the Bayesian Network Inference approach, and the other is the Granger Causality approach.

Bayesian networks are probabilistic graphical models initially introduced by [Kim & Pearl, 1987]. A Bayesian network is the specific type of graphical model which is directed acyclic graph. Causal influence measurement notation for time series was firstly proposed by Wiener-Granger. We can determine a causal influence of one time series on another, if the predication of one time series can be improved by incorporating the knowledge of the second one.

For the Bayesian network inference, it is often known to be applied to the static data. However, one can use dynamic Bayesian networks to deal with time series data. For the Granger causality approach, it is most likely to be applied to the time series data. It has the advantage to be able to decompose the data from time domain to frequency domain, so that one can clearly find out two variables interacting with each other at what frequency.

Giving the multi-variable time series data, the dynamic Bayesian networks and Granger causality can both be applied on it, but the causal relationship structures derived from these two approaches could be different. It is necessary to compare these two causal relationship inferring approaches in a general way. By doing the comparison, one expects to find the advantages, performances and stabilities for each technique. Firstly, we give an introduction on both Bayesian network and Granger causality used in this paper. In order to test these two approaches, we create a toy model and use Monte Carlos method to compare with each other for different situations, such as small sample size, random coefficients and so on.