Applied Causality Assignment 0

Feihan Lu

January 23, 2017

1 Problem

The problem I have is to develop a Vector Autoregressive Model or VAR to do inference on functional/effective connectivity among brain regions using resting-state multiple-subject-multiple-session fMRI time series data, i.e., there are more than one subjects who carried out no tasks during the scan, and each of them has more than one fMRI scans. The goal is to understand what brain regions affects others within how many time lags. In the literature so far, a variety of VAR models, with or without regularizing the coefficients, many of which in Bayesian frameworks, have been proposed for single- or multiple-subject data. However, none takes into account multiple sessions. Also, the number of brain regions considered so far has been small, such as 5 or 6. My goal is to build a model that can handle large number of subjects with multiple sessions, from which one can do inference on connectivity between large number of brain regions. This is a major focus of my thesis.

2 Data

The data I have are resting-state fMRI time series. The first dataset involves 800 subjects, each of whom has 4 repeat fMRI scans over 1200 time points, 2 seconds each, on 15 brain regions of interest, or ROIs. The second dataset contains 20 subjects with 2 repeat scans, over 200 time points on 39 ROIs. Both of them are of large scales.

3 Method

So far, I have developed a hierarchical sparse VAR model for small number of subjects, such as 5 to 10, with multiple sessions that penalizes population-level coefficients using multivariate Laplace-type prior, namely a double adaptive elastic-net prior proposed by [1]. Specifically, for subject $n = 1, \ldots, N$ and session $j = 1, \ldots, J$, I fit a usual VAR model with vectorized coefficients w_{nj}. Then, I suppose $w_{nj} = w + v_n + u_j$, where $v_n \sim MN(\mathbf{0}, \Omega_v^{-1}))$ is the subject-level random effects, and $u_j \sim MN(\mathbf{0}, \Omega_u^{-1})$ is the session-level random effects. Finally I use the double adaptive elastic-net prior for w, the population-level coefficients. I am using auxiliary Gibbs sampler to draw from posterior distribution. The problem with the current model is that it involves a large number of parameters, which is more than 10K, even for small number of subjects such as 5 to 10, and is prohibitive for large number of subjects, such as 800. I hope to develop a more scalable model along with faster computation. This is my goal of this course.

References

[1] Deborah Gefang. *Bayesian doubly adaptive elastic-net Lasso for VAR shrinkage*. International Jounal of Forecasting 30 (2014) 1-11.