Discussion of

Structural Learning of Contemporaneous Dependencies in Graphical VAR Models by Lucia Paci

Ed George
Wharton, University of Pennsylvania

O'Bayes 2019
University of Warwick
July 1, 2019

A Synopsis of the Contribution

- I) Overall goal: Build a VAR model $Y = ZB + E$ for q interrelated time series. Many, many parameters.
- II) Presence of lag coefficients and precision matrix parameters are indicated by the presence of edges in a sequence of evolving graphical models.
- III) Add priors on latent 0-1 indicators for every edge.
- IV) Use objective Fractional Bayes Factors to induce posterior probabilities on the submodels.
- V) Apply an MH strategy for posterior simulation.
- VI) Consider *HPM*, *MPM* and *FDRM* selection strategies.
- VII) Comparisons on Simulated and Real Data.

The Priors on the Edges

- Dynamic component

$$p(\boldsymbol{\Gamma}) \propto \binom{kq^2}{p^*} \Gamma\left(1+p^*\right) \Gamma\left(1+kq^2-p^*\right)$$

where $p^* = \sum_{h=1}^{kq^2} \mathrm{vec}(\boldsymbol{\Gamma})_h$.

- Contemporaneous component

$$p(G^u) \propto \binom{q(q-1)/2}{|G^u|} \Gamma\left(1+|G^u|\right) \Gamma\left(b+q(q-1)/2-|G^u|\right),$$

where $|G^u|$ denotes the number of edges in G^u and $b = (2q-2)/3 - 1$ to favor sparsity.

- Adaptive Beta Binomial priors which can encourage sparsity.
- Homogenous over models of the same size.
- May be interesting to consider structured elaborations.

Collapsed Gibbs Sampling of the Posterior

- Metropolis-Hastings proposal: At each step, locally modify $\mathbf{\Gamma}$ and G^u by adding and deleting edges one at a time.
- Could this one at a time strategy interfere with the required decomposability of G^u for likelihood factorization?
- Would it make sense to add and delete groups of edges?
- What about strategies such as switching?

The Fractional Bayes Factor

The FBF of model M_l against $M_{l'}$ is

$$\mathrm{FBF}_{ll'} = m_l^F(\boldsymbol{Y})/m_{l'}^F(\boldsymbol{Y}),$$

where $m_l^F(\boldsymbol{Y})$ is the **fractional marginal likelihood** of M_l given by

$$m_l^F(\boldsymbol{Y}) = \frac{\int f_l(\boldsymbol{Y} \mid \boldsymbol{\theta}_l) p_l^D(\boldsymbol{\theta}_l) d\boldsymbol{\theta}_l}{\int f_l^b(\boldsymbol{Y} \mid \boldsymbol{\theta}_l) p_l^D(\boldsymbol{\theta}_l) d\boldsymbol{\theta}_l} = \int f_l^{1-b}(\boldsymbol{Y} \mid \boldsymbol{\theta}_l) p_l^F(\boldsymbol{\theta}_l) d\boldsymbol{\theta}_l,$$

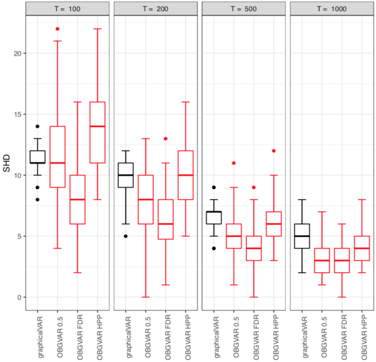
where $p_l^F(\boldsymbol{\theta}_k) \propto f_l^b(\boldsymbol{Y} \mid \boldsymbol{\theta}_l) p_l^D(\boldsymbol{\theta}_l)$ is the induced **fractional prior**.

- Starting with an improper noninformative prior, the fractional prior here becomes a *Matrix Normal Hyper-inverse Wishart distribution*. Tractable and proper!
- Is the tuning fraction $b = T_0/T$ used here arbitrary?
- Does it matter?

Three Selection Criteria Considered

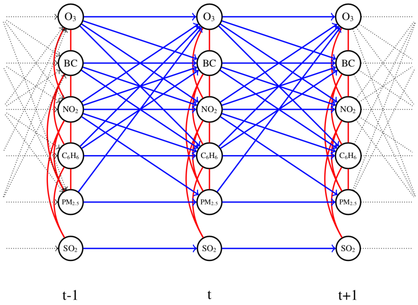
- *HPP* - Highest Posterior Probability Model.
- *MPM* - Median Posterior Probability Model. Include edges with marginal inclusion probability greater than 0.5.
- *FDRM* - False Discovery Rate Model. Include edges with marginal inclusion probability greater than r chosen such that $FDR \leq 0.05$.
- In canonical linear variable selection, *MPM* and *HPP* often select the same model, (though *MPM* is usually better when they differ). (Barbieri et al 2019).
- Because marginal inclusion probabilities are rapidly obtained via MCMC, *MPM* and *FDRM* are much easier to find than *HPP*.
- *MPM* is often the best single model approximation to the model average under quadratic loss.
- Is there a similar sense in which the *FDRM* is an approximation to some kind of optimality?

Simulation Comparisons for this VAR Problem



- Surprise - *FDRM* does better than *MPM*, which does better than *HPP*.
- Why does *FDRM* win here?
- L_1 penalized straw man here was beaten by Deshpande et al (2019). May be a good competitor here.

The Selected Air Quality Model



- Will be interesting to see what happens for k larger than 1, to see which higher order lagged effects appear.
- Would the *MPM* and *HPP* models be different here?
- Would other summaries of the selected model be useful?

Congratulations Lucia!