Annals of the Institute of Statistical Mathematics (2020) 72:33–36
https://doi.org/10.1007/s10463-019-00742-2

INVITED ARTICLE: SECOND AKAIKE MEMORIAL LECTURE

Discussion of "Bayesian forecasting of multivariate time series: scalability, structure uncertainty and decisions"

Jouchi Nakajima[1]

Received: 7 October 2019 / Published online: 15 December 2019

Abstract
The author focuses on the "decoupling and recoupling" idea that can critically increase both computational and forecasting efficiencies in practical problems for economic and financial data. My discussion is twofold. First, I briefly describe the idea with an example of time-varying vector autoregressions, which are widely used in the context. Second, I highlight the issue of how to assess patterns of simultaneous relationships.

Keywords Bayesian forecasting · Decouple/recouple · Time-varying vector autoregressions · Multivariate time-series models

1 Introduction

I thank the author for a great discussion of recent advances in Bayesian multivariate time-series modeling strategies with several relevant and practical examples in economics and financial data problems. I believe that his comprehensive description of key model structure and methods as well as notes on challenges and opportunities are all beneficial to readers. One of the main focuses in the paper is the decoupling and recoupling idea for estimating and forecasting multivariate time-series models. For high-dimensional problems, in particular, the idea is one of the strengths of the Bayesian approach. To review it, I briefly describe an example of time-varying vector autoregressions (TV-VAR) and see how the idea is applied to the model in a practi-

The Related Articles are https://doi.org/10.1007/s10463-019-00741-3; https://doi.org/10.1007/s10463-019-00743-1; https://doi.org/10.1007/s10463-019-00744-0.

The views expressed herein are those of the author alone and do not necessarily reflect those of the Bank of Japan.

✉ Jouchi Nakajima
jouchi.nakajima@boj.or.jp

[1] Bank of Japan, Chuo-ku, Tokyo 103-0021, Japan

cal setting. Then, I discuss the issue of simultaneous relationships that is one of the important aspects in the decoupling and recoupling strategy.

2 An example: time-varying vector autoregressions

The VAR models have been popular workhorses in macro- and financial econometrics, and the time-varying versions, TV-VAR models, have become quite popular since Primiceri (2005) developed a seminal form of the TV-VAR with stochastic volatility. Yet, the model structure itself was not new: it simply forms a traditional dynamic linear model (e.g., West and Harrison 1997). The Primiceri's model, specialized for an analysis with macroeconomic variables, fits a variety of contexts well, in particular, fiscal and monetary policy discussions (see also Nakajima 2011). In financial econometrics, Diebold and Yilmaz (2009) exploit the VAR model to assess spillover effects among financial variables such as stock price and exchange rates, and Geraci and Gnabo (2018) extend the framework with the TV-VAR.

Define a response $\boldsymbol{y}_t$, $(t = 1, 2, \ldots)$, as the $q \times 1$ vector. The TV-VAR(p) model forms

$$\boldsymbol{A}_t \boldsymbol{y}_t = \sum_{j=1}^{p} \boldsymbol{F}_{jt} \boldsymbol{y}_{t-j} + \boldsymbol{\varepsilon}_t, \quad \boldsymbol{\varepsilon}_t \sim N(\boldsymbol{0}, \boldsymbol{\Lambda}_t),$$

where $\boldsymbol{F}_{jt}$ is the $q \times q$ matrix of lag coefficients, and $\boldsymbol{\Lambda}_t$ is the $q \times q$ diagonal volatility matrix with ith diagonal element denoted by σ_{it}^2. Note that the model can include time-varying intercepts and regression components with other explanatory variables, although these additional ingredients do not change the following discussion.

The $\boldsymbol{A}_t$ is the $q \times q$ matrix that defines simultaneous relationship among q variables, which is analogous to simultaneous parents and *parental predictors* in the author's discussion. With the diagonal structure of $\boldsymbol{\Lambda}_t$, the $\boldsymbol{A}_t$ defines patterns of contemporaneous dependencies among the responses $\{y_{1t}, \ldots, y_{qt}\}$. For identification, the model requires at least $q(q-1)/2$ elements in the off-diagonal part of $\boldsymbol{A}_t$ set to be zero.

A typical assumption for the contemporaneous structure in macroeconomic and financial variable data contexts is a triangular matrix:

$$\boldsymbol{A}_t = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ -a_{21t} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ -a_{q1t} & \cdots & -a_{q,q-1,t} & 1 \end{pmatrix}.$$

This leads to an implied reduced model form:

$$\boldsymbol{y}_t = \sum_{j=1}^{p} \boldsymbol{B}_{jt} \boldsymbol{y}_{t-j} + \boldsymbol{\nu}_t, \quad \boldsymbol{\nu}_t \sim N(\boldsymbol{0}, \boldsymbol{\Sigma}_t), \tag{1}$$

where $\boldsymbol{B}_{jt} = \boldsymbol{A}_t^{-1}\boldsymbol{F}_{jt}$, for $j = 1 : p$, and $\boldsymbol{\Sigma}_t = \boldsymbol{A}_t^{-1}\boldsymbol{\Lambda}_t\boldsymbol{A}_t'^{-1}$. We can see that the variance matrix of the innovation, $\boldsymbol{\Sigma}_t$, forms a Cholesky-style decomposition with $\boldsymbol{A}_t$ and $\boldsymbol{\Lambda}_t$. This restricts $q(q-1)/2$ elements in $\boldsymbol{A}_t$ to be zero, and so requires no additional constraints for identification. The parental predictors of DDNMs (in Section 3) have the same structure as the contemporaneous relationship relies on only one side (upper or lower) of the triangular part in $\boldsymbol{A}_t$. The discussion of the DDNMs assumes more sparse structure as q increases, i.e., most of a_{ijt}'s are potentially zero.

A decoupling step is implemented by recasting the model as a triangular set of univariate dynamic regressions:

$$\begin{aligned}
y_{1t} &= \boldsymbol{b}_{1t}'\boldsymbol{x}_{t-1} + \varepsilon_{1t},\\
y_{2t} &= a_{21t}y_{1t} + \boldsymbol{b}_{2t}'\boldsymbol{x}_{t-1} + \varepsilon_{2t},\\
y_{3t} &= a_{31t}y_{1t} + a_{32t}y_{2t} + \boldsymbol{b}_{3t}'\boldsymbol{x}_{t-1} + \varepsilon_{3t},\\
&\vdots \qquad \vdots\\
y_{qt} &= a_{q1t}y_{1t} + \cdots + a_{q,q-1,t}y_{q-1,t} + \mathbf{b}_{qt}'\mathbf{x}_{t-1} + \varepsilon_{qt},
\end{aligned}$$

where $\boldsymbol{x}_{t-1}$ is the $pq \times 1$ vector of lagged responses, defined by $\boldsymbol{x}_{t-1}' = (\boldsymbol{y}_{t-1}', \ldots, \boldsymbol{y}_{t-p}')$; $\boldsymbol{b}_{it}$ is the corresponding vector that consists of lag coefficient elements in $\boldsymbol{B}_{jt}$'s; and $\varepsilon_{it} \sim N(0, \sigma_{it}^2)$, for $i = 1 : q$. The key technical benefit is $\mathrm{Cov}(\varepsilon_{it}, \varepsilon_{js}) = 0$, for $i \neq j$ as well as for all t, s. Under conditionally independent priors over the coefficient processes and parameters, the model structure enables us to estimate q univariate dynamic regression models separately, and in parallel. Gains in computational efficiency are relevant, in particular as q increases, i.e., in higher-dimensional problems.

Then, posterior estimates from the decoupling step are fed into the recoupling step for forecasting and decisions. The recoupled model is basically based on Eq. (1), where the $\boldsymbol{A}_t$ elements link ("cross-talk") contemporaneous relationships among the y_{it}. Sequential forecasting and intervention analyses are straightforward with the reduced form equations.

3 Contemporaneous relationship

As discussed by the author in the paper, the ordering of the responses in $\boldsymbol{y}_t$, and more generally, the structure of $\boldsymbol{A}_t$ can be the issue. As far as an interest in forecasting is concerned, ordering is almost irrelevant because a predictive distribution relies only on the resulting covariance matrix $\boldsymbol{\Sigma}_t$ in Eq. (1). However, some other analysis such as intervention and impulse response analysis may suffer from the issue.

There are mainly three formal approaches to address the structure of $\boldsymbol{A}_t$. One way is a use of economic theory or "prior" based on economic reasonings. In macroeconomics, the Cholesky-style decomposition has been widely used with the ordering determined based on some economic reasoning (Sims 1980). For example, the interest rate is often placed last in the ordering as changes in the interest rate reflect contemporaneous changes in other macroeconomic variables such as output and inflation rate. Christiano et al. (1999) propose a block recursive approach that restricts several elements in the triangular part to be zero.

The second approach is based on model fit and forecasting performance: one example is described in the SGDLM application (in Section 4.6). This gives an “optimal” pattern of the simultaneous parents in terms of forecasting, while some priors or constraints may be required if q is quite large. The example in the paper sets $|pa(j)| = 20$, for $q = 401$, assuming relatively few series have conditional contemporaneous relationships with others. The third approach is a full analysis, searching for the best patterns of the simultaneous parents over all the possible combinations. When q is small, it is possible to implement even if $\boldsymbol{A}_t$ is time varying (see e.g., Nakajima and West 2013, 2015). However, if q is large, it would be almost infeasible due to the computational burden in practice. Finally, a mixture of the theory-based approach and more data-based approaches could be suitable depending on data and context.

References

Christiano, L. J., Eichenbaum, M., Evans, C. L. (1999). Monetary policy shocks: What have we learned and to what end? In J. B. Taylor & M. Woodford (Eds.), *Handbook of macroeconomics*, Vol. 1A, pp. 65–148. Amsterdam: Elsevier Science.

Diebold, F. X., Yilmaz, K. (2009). Measuring financial asset return and volatility spillovers, with application to global equity markets. *Economic Journal*, *119*, 158–171.

Geraci, M. V., Gnabo, J. Y. (2018). Measuring interconnectedness between financial institutions with Bayesian time-varying vector autoregressions. *Journal of Financial and Quantitative Analysis*, *53*, 1371–1390.

Nakajima, J. (2011). Time-varying parameter VAR model with stochastic volatility: An overview of methodology and empirical applications. *Monetary and Economic Studies*, *29*, 107–142.

Nakajima, J., West, M. (2013). Bayesian analysis of latent threshold dynamic models. *Journal of Business and Economic Statistics*, *31*, 151–164.

Nakajima, J., West, M. (2015). Dynamic network signal processing using latent threshold models. *Digital Signal Processing*, *47*, 5–16.

Primiceri, G. E. (2005). Time varying structural vector autoregressions and monetary policy. *Review of Economic Studies*, *72*, 821–852.

Sims, C. A. (1980). Macroeconomics and reality. *Econometrica*, *48*, 1–48.

West, M., Harrison, P. J. (1997). *Bayesian forecasting and dynamic models*, 2nd edn. New York: Springer.