Annals of the Institute of Statistical Mathematics (2020) 72:37–39
https://doi.org/10.1007/s10463-019-00743-1

INVITED ARTICLE: SECOND AKAIKE MEMORIAL LECTURE

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

Chris Glynn[1]

Received: 9 October 2019 / Published online: 9 December 2019
© The Institute of Statistical Mathematics, Tokyo 2019

I congratulate Professor West for his 2018 Akaike Memorial Lecture Award and for articulately synthesizing recent research in this unified treatment of the “decouple/recouple” framework. For readers who learned Bayesian dynamic models from West and Harrison (1997), the motivation and multivariate extensions of univariate dynamic linear models (DLMs) are familiar. The current focus on modeling sparse cross-series structure for scaling posterior computations to high dimensions is a welcome addendum.

A current challenge in Bayesian analysis is to scale models and computational procedures to meet the demands of increasingly large and complex data without sacrificing fundamentals of applied statistics. Physical, social, and economic sciences rely heavily on statistical models that are richly structured, interpretable, and reliably quantify uncertainty. There is great value to science in models that are both interpretable *and* scalable. In this regard, the decouple/recouple framework is an important contribution for modeling large collections of time series.

The computational gains of the decouple/recouple framework are achieved by exploiting sparse cross-series structure. At each time t, it is assumed that series j has a known set of simultaneous parents, denoted $sp(j) \subseteq \{1 : q\}\backslash\{j\}$. In the SGDLM setting, the observation of series j at time t is modeled as the composition of regressions on series-specific covariates $\boldsymbol{x}_{j,t}$ and simultaneous parental series $\boldsymbol{y}_{sp(j),t}$,

$$y_{j,t} = \boldsymbol{x}'_{j,t}\phi_{j,t} + \boldsymbol{y}'_{sp(j),t}\gamma_{j,t} + \nu_{j,t}. \tag{1}$$

The $\gamma_{j,t}$ state vector is augmented by zeros to form the jth row in matrix $\boldsymbol{\Gamma}_t$, which encodes the joint collection of simultaneous parent relationships across all series at time t. In $\boldsymbol{\Gamma}_t$, the jkth element is nonzero if $y_{k,t}$ is a parent of $y_{j,t}$, and the analysis

The Related Articles are https://doi.org/10.1007/s10463-019-00741-3; https://doi.org/10.1007/s10463-019-00742-2; https://doi.org/10.1007/s10463-019-00744-0.

✉ Chris Glynn
christopher.glynn@unh.edu

1 Paul College of Business and Economics, University of New Hampshire, Durham, NH, USA

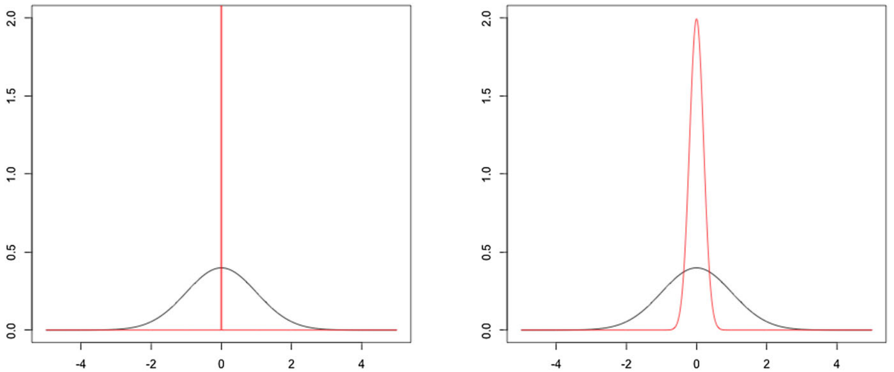

Fig. 1 Mixture prior distribution with a $N(0,1)$ component a point mass at zero (left) and spike and slab prior (right)

assumes independent Gaussian prior distributions for each nonzero coefficient. The prior distribution is then

$$\gamma_{j,k,t} \sim N(m_\gamma, \sigma_\gamma^2)\mathbb{1}_{\{k \in sp(j)\}} + \left(1 - \mathbb{1}_{\{k \in sp(j)\}}\right)\delta_0(\gamma_{j,k,t}), \text{ when } j \neq k. \quad (2)$$

While practical computational considerations in the SGDLM framework require that the set of simultaneous parents is either known or estimated with a heuristic algorithm (Gruber and West 2017), modeling the probability that series k is included in $sp(j)$ sheds light on connections between the decouple/recouple framework and other well-known variable selection methods. In addition, it points to interesting directions of future research. Suppose a model extension where $P(k \in sp(j)) = \pi$ is the prior probability that series k is in $sp(j)$. Then, the prior distribution for state variable $\gamma_{j,k,t}$ would be a mixture

$$\gamma_{j,k,t} \sim \pi N(m_\gamma, \sigma_\gamma^2) + (1 - \pi)\delta_0(\gamma_{j,k,t}) \quad (3)$$

where one component is the standard $N(m_\gamma, \sigma_\gamma^2)$ prior when $k \in sp(j)$ and the other is a point mass at zero when $k \notin sp(j)$. This scenario is illustrated in Fig. 1 (left).

An interesting direction of future research is to relax the assumption that $\gamma_{j,k,t} = 0$ when $k \notin sp(j)$ and introduce a Gaussian noise centered at zero instead. The exact zero that encodes sparse structure is elegant; however, there may be further computational gains to be achieved by allowing contributions from series that are *approximately* rather than exactly zero. In this spike and slab type setting (George and McCulloch 1993), the prior for each $\gamma_{j,k,t}$ is a mixture of two Gaussian distributions, the original Gaussian component and a Gaussian component tightly concentrated around zero [Fig. 1 (right)]. When utilizing simultaneous values from other time series $\boldsymbol{y}_{sp(j),t}$ as regressors, choosing which series to include in $sp(j)$ is a dynamic variable selection problem, and Rockova and McAlinn (2017) utilize the spike and slab prior to model dynamic sparse structure.

While Professor West makes it clear that, in the present work, the use of graphical structure is a means to improve forecast performance in multivariate time series and not a key inference goal, there are applications where inferring parental relationships in cross-series structure is important. One example is managing the risk of contagion in financial crises. Given a large collection of real-time stock price data for systemically important financial institutions, inferring simultaneous parents of individual institutions (Wells Fargo, for example) is useful to both regulators and policymakers. Learning simultaneous parents is especially important when designing market interventions to prevent (or halt) contagion. Rather than making investments in all systemically important banks, as the Federal Reserve did in the financial crisis of 2008, a central bank could make targeted investments in the few firms that are simultaneous parents to many other institutions.

References

George, E. I., McCulloch, R. E. (1993). Variable selection via Gibbs sampling. *Journal of the American Statistical Association*, *88*(423), 881–889.

Gruber, L. F., West, M. (2017). Bayesian forecasting and scalable multivariate volatility analysis using simultaneous graphical dynamic linear models. *Econometrics and Statistics*, *3*, 3–22. arXiv:1606.08291.

Rockova, V., McAlinn, K. (2017). Dynamic variable selection with spike-and-slab process priors.

West, M., Harrison, J. (1997). *Bayesian Forecasting and Dynamic Modeling*2nd ed. New York, NY: Springer.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.