International Journal of Forecasting 33 (2017) 627–651

Contents lists available at ScienceDirect

International Journal of Forecasting

journal homepage: www.elsevier.com/locate/ijforecast

VARX-L: Structured regularization for large vector autoregressions with exogenous variables

William B. Nicholson [a,*], David S. Matteson [b,c], Jacob Bien [d,e]

[a] Department of Statistical Science, Cornell University, 301 Malott Hall, Ithaca, NY 14853, United States
[b] Department of Statistical Science, Cornell University, 1196 Comstock Hall, Ithaca, NY 14853, United States
[c] Department of Social Statistics, Cornell University, 1196 Comstock Hall, Ithaca, NY 14853, United States
[d] Department of Biological Statistics and Computational Biology, Cornell University, 1178 Comstock Hall, Ithaca, NY 14853, United States
[e] Department of Statistical Science, Cornell University, 1178 Comstock Hall, Ithaca, NY 14853, United States

ARTICLE INFO

Keywords:
Big data
Forecasting
Group lasso
Macroeconometrics
Time series

ABSTRACT

The vector autoregression (VAR) has long proven to be an effective method for modeling the joint dynamics of macroeconomic time series, as well as for forecasting. One major shortcoming of the VAR that has limited its applicability is its heavy parameterization: the parameter space grows quadratically with the number of series included, quickly exhausting the available degrees of freedom. Consequently, using VARs for forecasting is intractable for low-frequency, high-dimensional macroeconomic data. However, empirical evidence suggests that VARs that incorporate more component series tend to result in more accurate forecasts. Most conventional methods that allow for the estimation of large VARs either require *ad hoc* subjective specifications or are computationally infeasible. Moreover, as global economies become more intricately intertwined, there has been a substantial interest in incorporating the impact of stochastic, unmodeled *exogenous* variables. Vector autoregression with exogenous variables (VARX) extends the VAR to allow for the inclusion of unmodeled variables, but faces similar dimensionality challenges.

This paper introduces the VARX-L framework, a structured family of VARX models, and provides a methodology that allows for both efficient estimation and accurate forecasting in high-dimensional analysis. VARX-L adapts several prominent scalar regression regularization techniques to a vector time series context, which greatly reduces the parameter space of VAR and VARX models. We also highlight a compelling extension that allows for shrinking toward reference models, such as a vector random walk. We demonstrate the efficacy of VARX-L in both low- and high-dimensional macroeconomic forecasting applications and simulated data examples. Our methodology is easy to reproduce in a publicly available R package.
© 2017 International Institute of Forecasters. Published by Elsevier B.V. All rights reserved.

1. Introduction

The practice of macroeconomic forecasting was spearheaded by Klein and Goldberger (1955), whose eponymous simultaneous equation system forecast the behaviors of 15 annual macroeconomic indicators jointly, including

* Corresponding author.
E-mail addresses: wbn8@cornell.edu (W.B. Nicholson), matteson@cornell.edu (D.S. Matteson), jbien@cornell.edu (J. Bien).
URLs: http://www.wbnicholson.com (W.B. Nicholson), http://www.stat.cornell.edu/~matteson/ (D.S. Matteson), http://faculty.bscb.cornell.edu/~bien/ (J. Bien).

http://dx.doi.org/10.1016/j.ijforecast.2017.01.003
0169-2070/© 2017 International Institute of Forecasters. Published by Elsevier B.V. All rights reserved.

consumer expenditures, interest rates, and corporate profits. The parameterization and identification restrictions of these models were influenced heavily by Keynesian economic theory. As computing power increased, such models became larger and began to utilize higher frequency data. Forecasts and simulations from these models were commonly used to inform government policymakers regarding the overall state of the economy and to influence policy decisions (Welfe, 2013). As the Klein-Goldberger model and its extensions were motivated primarily by Keynesian economic theory, the collapse of the Bretton Woods monetary system and severe oil price shocks led to widespread forecasting failures in the 1970s (Diebold, 1998). At this time, the vector autoregression (VAR), popularized by Sims (1980), emerged as an atheoretical forecasting technique that was underpinned by statistical methodology and not subject to the ebb and flow of contemporary macroeconomic theory.

Unfortunately, the flexibility of the VAR can create modeling complications. In the absence of prior information, the VAR assumes that every series interacts linearly with both its own past values and those of every other series included. Such a model is known as an *unrestricted* VAR. As most economic series are low-frequency (monthly, quarterly, or annual) there is rarely enough data available to allow accurate forecasts using large unrestricted VARs. Such models are overparameterized, provide inaccurate forecasts, and are very sensitive to changes in economic variables. Consequently, the VAR's parameter space must be reduced in such applications, either in a data-driven manner or based upon the modeler's knowledge of the underlying economic system. This model selection process has been described as "blending data and personal beliefs according to a subjective, undocumented procedure that others cannot duplicate" (Todd, 1990, p. 18).

Despite their overparameterization, large VARs can be preferable to their smaller counterparts in many applications, as small models can exclude potentially relevant variables. Ideally, a variable should always be included in the model unless one has prior knowledge that it is irrelevant. For example, as was described by Lütkepohl (2014), modeling the Taylor rule (Taylor, 1993) requires an estimate of the "output gap" between the real gross domestic product and the potential output. The output gap is difficult to measure and can include many explanatory variables that encompass disaggregated economic measurements. Moreover, recent work by Hendry and Hubrich (2011) and Ibarra (2012) showed that incorporating disaggregated series improves the forecasts of macroeconomic aggregates such as the consumer price index. Hence, a large vector autoregression with a coherent variable selection procedure is required in order to utilize all relevant economic information properly in these scenarios.

Shortly after the inception of the VAR, efforts were made to develop a systematic approach to the reduction of its parameterization. Early attempts, such as that of Litterman (1979), centered upon a Bayesian approach underpinned by contemporary macroeconomic theory. In applying a Bayesian VAR with a Gaussian prior (analogous to ridge regression), specific priors were formulated based upon stylized facts regarding US macroeconomic data. For example, the popular *Minnesota prior* incorporates the prevailing belief that macroeconomic variables can be modeled reasonably by means of a univariate random walk, by shrinking the estimated models toward univariate unit root processes.

The Bayesian VAR with a Minnesota prior was shown by Robertson and Tallman (1999) to produce forecasts that were superior to those from the conventional VAR, univariate models, and traditional simultaneous equation models. However, this approach is very restrictive; in particular, it assumes that all series are contemporaneously uncorrelated, and requires the specification of several hyperparameters. Moreover, the Minnesota prior cannot accommodate large VARs itself. As was pointed out by Banbura, Giannone, and Reichlin (2009), when constructing a 40-variable system, in addition to the Minnesota prior, Litterman (1986b) also imposed strict economically-motivated restrictions in order to limit the number of variables in each equation.

The modern Bayesian VAR extensions that were originally proposed by Kadiyala and Karlsson (1997) and compiled by Koop (2013) show that empirical regularization methods alone allow for the accurate forecasting of large VARs. Such procedures impose data-driven restrictions on the parameter space while allowing for more general covariance specifications and the estimation of hyperparameters via empirical Bayes or Markov chain Monte Carlo methods. These approaches are computationally expensive, and multi-step forecasts are nonlinear and must be obtained by additional simulation. Using a conjugate Gaussian-Wishart prior, Banbura et al. (2009) extended the Minnesota prior to a high-dimensional setting with a closed-form posterior distribution. Their technique uses a single hyperparameter, which is estimated by cross-validation. However, their specification does not perform variable selection, and their penalty parameter selection procedure is more natural within a frequentist framework.

More recent attempts to reduce the parameter space of VARs have incorporated the *lasso* (Tibshirani, 1996), a least squares variable selection technique. These approaches include the *lasso-VAR* that was proposed by Hsu, Hung, and Chang (2008) and explored further by Davis, Zang, and Zheng (2016), Li and Chen (2014), and Song and Bickel (2011). Its theoretical properties were investigated by Basu and Michailidis (2015) and Kock and Callot (2015). Gefang (2012) considered a Bayesian implementation of the elastic net, an extension of the lasso proposed by Zou and Hastie (2005) that accounts for highly correlated covariates. However, their implementation is not computationally tractable and little forecasting improvement relative to existing methods is observed. The lasso-VAR has several advantages over Bayesian approaches, as it is more computationally efficient in high dimensions, performs variable selection, and can readily compute multi-step forecasts and their associated prediction intervals.

In many applications, a VAR's forecasts can be improved through the incorporation of unmodeled exogenous variables, which are determined outside the VAR. Examples of exogenous variables are context-dependent and range from leading indicators to weather-related measurements. In many scenarios, global macroeconomic variables, such

Download English Version:

https://daneshyari.com/en/article/5106424

Download Persian Version:

https://daneshyari.com/article/5106424

Daneshyari.com