Prior Selection in Bayesian Vector Autoregressive Models

Domenico Giannone (ULB)
Michele Lenza (ECB)
Giorgio Primiceri (Northwestern)

Vector autoregressions (VARs) are flexible time series models that can capture complex dynamic interrelationships among macroeconomic variables. Their generality, however, comes at the cost of being very densely parameterized. As a result, the estimation of VARs tends to deliver good in-sample fit, but unstable inference and inaccurate out-of-sample forecasts, particularly when the model includes many variables.

A potential solution to this problem is combining the richly parameterized unrestricted model with parsimonious priors, which help controlling estimation uncertainty (see, for example, Doan, Litterman and Sims, 1982 and Litterman, 1986). In fact, De Mol, Giannone and Reichlin (2008) and Banbura, Giannone and Reichlin (2009) have shown that Bayesian VARs are able to provide stable and reliable predictions even with very high dimensional data.

Unfortunately, however, the issue of how to optimally set the weight of the prior relative to the likelihood information is largely unexplored and has mostly relied on ad hoc procedures. For example, Doan, Litterman and Sims (1982) selected the weight of the prior model by optimizing the out-of-sample forecasting performances of their model. In subsequent work the choice of priors has often been subjective or it has been closely linked to Doan, Litterman and Sims (1982), despite the heterogeneity of research questions and datasets. The perceived arbitrariness of the prior choice is certainly one reason why the potential of Bayesian VARs has not been fully exploited.

In this paper, we provide a simple and theoretically founded methodology for prior selection in Bayesian VARs. Our recommendation is to select priors by maximizing (or integrating over) the likelihood function integrated over the model parameters. The latter is known as the “marginal” data density and it only depends on the hyper-parameters that characterize the relative weight of the prior model and the information in the data. Our suggested procedure has several appealing features:

1. It is the most theoretically clean Bayesian procedure for data driven prior selection
2. Its solution can be interpreted as the result of maximizing the model’s out-of-sample forecasting accuracy. Moreover, we explore the direct connection with the classical literature on random effects and the Bayesian literature on hierarchical models.
3. It is easy to perform since the marginal data density of Bayesian VARs (with conjugate priors) is available in closed form.

We present 2 applications of our methodology:

1. **Reduced-form analysis**: We show that the out-of-sample forecasting accuracy of our model is superior to that of VARs with flat priors and BVARs with ad hoc prior selection (especially when the curse of dimensionality becomes a serious problem) and comparable to that of factor models.

2. **Structural analysis**: We show that our model approximates well the solution of Dynamic Stochastic General Equilibrium models with very persistent MA components. This is because our procedure allows us to efficiently handle VARs with a large number of lags. As a consequence, our BVARs can be used to back out structural shocks and construct impulse responses, even when standard methods are fragile. We demonstrate this feature or our procedure by performing structural VAR analysis on artificial data generated from a DSGE model known to have a nearly non-invertible MA representation.