Bayesian analysis of VARs, State-space models and DSGEs Part IV - BVARs and State-Space Models

Mattias Villani

Division of Statistics and Machine Learning
Department of Computer and Information Science
Linköping University

Lecture overview

- **BVARs**
- **State-space models and the Kalman filter**

VARs

- **Vector Autoregressive Process (VAR)**

$$\underset{p\times 1}{x_t} = \sum_{k=1}^{K} \underset{p\times p}{\Pi_k} x_{t-k} + \underset{p\times q}{\Phi}\underset{q\times 1}{d_t} + u_t, \qquad u_t \overset{iid}{\sim} N(0,\Sigma)$$

- **Structural form**

$$\Upsilon x_t = \sum_{k=1}^{K} \Pi_k x_{t-k} + \Phi d_t + \varepsilon_t, \qquad u_t \overset{iid}{\sim} N(0, I_p)$$

- **Steady-state VAR [1]**

$$x_t = \Psi d_t + \sum_{k=1}^{K} \Pi_k \left(x_{t-k} - \Psi d_{t-k}\right) + u_t, \qquad u_t \overset{iid}{\sim} N(0,\Sigma)$$

  such that

$$Ex_t = \Psi d_t$$

- Important: **forecasts at long horizons** end up at the steady state!
- Steady-state VAR in structural/cointegration form is possible **[1]**.

Bayesian VARs

- **Steady-state VAR [1]**

$$x_t = \Psi d_t + \sum_{k=1}^{K} \Pi_k \left(x_{t-k} - \Psi d_{t-k} \right) + u_t, \qquad u_t \overset{iid}{\sim} N(0, \Sigma)$$

- Prior $p(\Pi, \Psi, \Sigma) = p(\Pi)p(\Psi)p(\Sigma)$ (prior independence)
  - $\mathrm{vec}(\Pi) \sim \mathit{MultivariateNormal}$
  - $\mathrm{vec}(\Psi) \sim \mathit{MultivariateNormal}$
  - Σ has a noninformative "uniform" prior.
- **Prior mean** is the univariate AR(1) processes:

$$\begin{aligned} x_{1t} &= \mu_{\Psi,1} + \mu_{\Pi,1} x_{1,t-1} + \varepsilon_{1t} \\ &\vdots \\ x_{pt} &= \mu_{\Psi,p} + \mu_{\Pi,p} x_{p,t-1} + \varepsilon_{pt} \end{aligned}$$

Bayesian VARs

- Steady-state VAR

$$x_t = \Psi d_t + \sum_{k=1}^{K} \Pi_k \left(x_{t-k} - \Psi d_{t-k}\right) + u_t, \qquad u_t \overset{iid}{\sim} N(0, \Sigma)$$

- **Minnesota prior standard deviation** for the elements in $\Pi_1, ..., \Pi_K$:

$$\mathrm{Std}(\pi_{ij}^{(k)}) = \begin{cases} \frac{\lambda_1 s_i}{k^{\lambda_3} s_j} & \textbf{if } i = j \textbf{ (own lag)} \\ \frac{\lambda_1 \lambda_2 s_i}{k^{\lambda_3} s_j} & \textbf{if } i \neq j \textbf{ (cross-equation lag)} \end{cases}$$

  where s_i is the residual standard deviation from OLS fit of AR(K) process to ith time series.

- **Prior on steady state** for variable j:

$$N\left(\mu_{\Psi,j}, \tau_j^2\right)$$

  specified by an expert or a survey among experts.

Gibbs sampling for Bayesian VARs

- The **joint posterior distribution** $p(\Pi, \Psi, \Sigma | \mathbf{x}_{1:T})$ can be obtained by **Gibbs sampling**[1]
  - Simulate $\Pi | \Psi, \Sigma, \mathbf{x}_{1:T}$ from a multivariate Normal
  - Simulate $\Psi | \Pi, \Sigma, \mathbf{x}_{1:T}$ from a multivariate Normal
  - Simulate $\Sigma | \Pi, \Psi, \mathbf{x}_{1:T}$ from a inverse Wishart
- Use OLS/ML as initial values.
- **MCMC convergence** is much less of a problem than for DSGEs.
- Exception: weak prior for Ψ in a very persistent process (but why use the steady-state formulation then?)
- The **joint predictive distribution** $p(\mathbf{x}_{T+1:t+h} | \mathbf{x}_{1:T})$ is obtained by simulation (see the AR model in part I).
- **Joint posterior distribution of the impulse responses** are obtained by simple computations from the Gibbs samples (see Part III).

BVARs - application to Swedish macro data

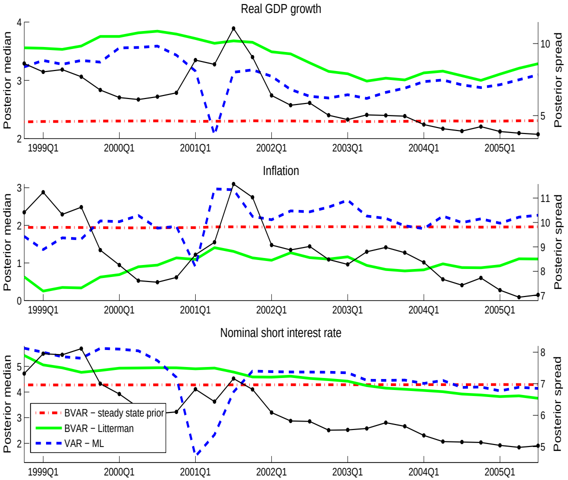
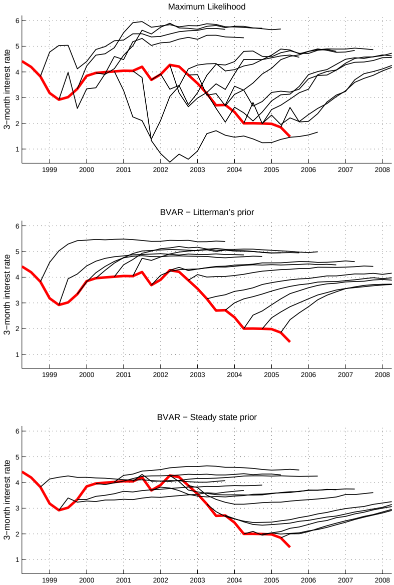
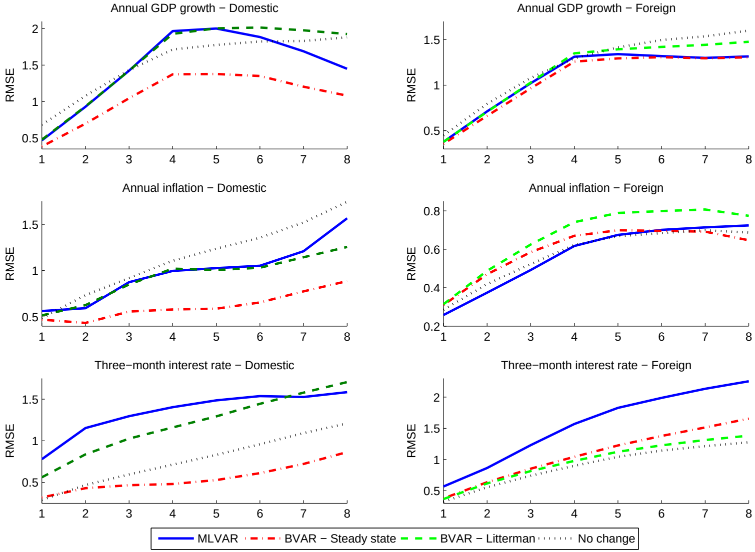


FIGURE 5. Swedish macro data. Sequential posterior median estimates of the steady state over time (measured on the left axis). The black solid line with dots (measured on the right axis) displays the length of the 95% probability interval of the steady state under the Litterman prior.

BVARs - application to Swedish macro data



Gibbs sampling for Bayesian VARs



State-space models

- **Observed (measured) variables** (GDP, CPI, interest rate etc) are driven by **unobserved latent variables** (e.g. technology or preference shocks).
- **State-space model**

$$\xi_t = \mathbf{F}\xi_{t-1} + \mathbf{v}_t, \quad \mathbf{v}_t \overset{iid}{\sim} N(0, \mathbf{Q}) \quad \text{(state transition equation)}$$
$$\mathbf{y}_t = \mathbf{H}'\xi_t + \mathbf{w}_t, \quad \mathbf{w}_t \overset{iid}{\sim} N(0, \mathbf{R}) \quad \text{(measurement equation)}$$

- Model parameters:
  - F, H, Q, R and $\xi_{1:T}$

Filtering

- Assume that F, H, Q, R are known.
- **Filtering distribution**

$$p(\xi_t | \mathbf{y}_{1:t}) = N(\xi_{t|t}, P_{t|t}).$$

- **Filtering distribution**: posterior distribution of the state at time $t$using data up to time t.
- Filtering distribution at all time periods ($\xi_{t|t}$ and $P_{t|t}$ for $t = 1 : T$) is obtained by iterating the **Kalman filter** forward. See [2].

Filtering = state marginalization

- Assume $\mathbf{F}, \mathbf{H}, \mathbf{Q}, \mathbf{R}$ are unknown. Wanted: **parameter posterior**

$$p(\mathbf{F}, \mathbf{H}, \mathbf{Q}, \mathbf{R} | \mathbf{y}_{1:T}) \propto p(\mathbf{y}_{1:T} | \mathbf{F}, \mathbf{H}, \mathbf{Q}, \mathbf{R}) p(\mathbf{F}, \mathbf{H}, \mathbf{Q}, \mathbf{R})$$

- Need to compute $p(\mathbf{y}_{1:T} | \mathbf{F}, \mathbf{H}, \mathbf{Q}, \mathbf{R})$ with $\xi_{1:T}$ marginalized out.
- **State marginalization** is exactly what the **Kalman filter** does

$$p(\mathbf{y}_{1:T} | \mathbf{F}, \mathbf{H}, \mathbf{Q}, \mathbf{R}) = \prod_{t=1}^{T} p(\mathbf{y}_t | y_{1:(t-1)}, \mathbf{F}, \mathbf{H}, \mathbf{Q}, \mathbf{R})$$

  where

$$p(\mathbf{y}_t | y_{1:(t-1)}, \mathbf{F}, \mathbf{H}, \mathbf{Q}, \mathbf{R}) = N\left(\mathbf{H}\xi_{t|t-1}, \mathbf{H}'\mathbf{P}_{t|t-1}\mathbf{H} + \mathbf{R}\right)$$

  and $\xi_{t|t-1}$ and $\mathbf{P}_{t|t-1}$ are by-products of the Kalman filter [2].
- In **DSGEs** we work with the likelihood for the 'deep' parameters θ

$$p(\mathbf{y}_{1:T} | \mathbf{F}(\theta), \mathbf{H}(\theta), \mathbf{Q}(\theta), \mathbf{R}(\theta))$$

Smoothing

- **Filtering distribution**

$$p(\xi_t|\mathbf{y}_{1:t}) = N(\xi_{t|t}, P_{t|t})$$

- **Smoothing distribution (state posterior)**:

$$p(\xi_t|\mathbf{y}_{1:T}) = N(\xi_{t|T}, P_{t|T})$$

  posterior distribution of the state at time t using data up to time t.

- Smoothing distribution at all time periods ($\xi_{t|T}$ and $P_{t|T}$ for $t = 1 : T$) is obtained by iterating the **Kalman filter backwards** starting from $\xi_{T|T}$ and $P_{T|T}$ obtained at the end of Kalman filtering.

- M. Villani, "Steady-state priors for vector autoregressions," *Journal of Applied Econometrics*, vol. 24, no. 4, pp. 630–650, 2009.
- J. D. Hamilton, *Time series analysis*, vol. 2.
  1994.