



Adventures in Bayesian Structural Time Series

Part 3: Structural Time Series in Bayesian Context

Andrew Bates, Josh Gloyd, Tyler Tucker



⊠ Without Regression:



⊠ Without Regression:

⊠ Prior



⊠ Without Regression:

- ⊠ Prior

- ⊠ Posterior



⊠ Without Regression:

⊠ Prior

⊠ Posterior

⊠ With Regression:



⊠ Without Regression:

⊠ Prior

⊠ Posterior

⊠ With Regression:

⊠ Prior



⊠ Without Regression:

- ⊠ Prior
- ⊠ Posterior

⊠ With Regression:

- ⊠ Prior
- ⊠ Posterior



⊠ Prior



⊠ Prior

- ⊠ Specify distribution on parameters (variances)



- ⊠ Prior

- ⊠ Specify distribution on parameters (variances)

- ⊠ Posterior



- ❖ Prior

- ❖ Specify distribution on parameters (variances)

- ❖ Posterior

- ❖ Kalman filter and Kalman smoother



Without Regression

- ⊠ Prior

- ⊠ Specify distribution on parameters (variances)

- ⊠ Posterior

- ⊠ Kalman filter and Kalman smoother
 - ⊠ Markov Chain Monte Carlo



- ⊗ Consider regression coefficients fixed through time



- ⊗ Consider regression coefficients fixed through time
 - ⊗ If not, include as state variables



- ⊠ Consider regression coefficients fixed through time
 - ⊠ If not, include as state variables
- ⊠ **Spike:** prior for probability of regression coefficients being zero



- ⊠ Consider regression coefficients fixed through time
 - ⊠ If not, include as state variables
- ⊠ **Spike:** prior for probability of regression coefficients being zero
- ⊠ **Slab:** prior on nonzero regression coefficients and variance



Spike and Slab Prior

Spike

- ⊠ $\gamma_k = 1$ if $\beta_k \neq 0$. $\gamma_k = 0$ if $\beta_k = 0$
- ⊠ β_γ : subset of β s with $\beta_k \neq 0$
- ⊠ Independent Bernoulli prior:

$$\gamma \sim \prod_{k=1}^K \pi_k^{\gamma_k} (1 - \pi_k)^{1-\gamma_k}$$



Spike and Slab Prior

Slab

- ⊠ Ω^{-1} : symmetric matrix
- ⊠ Ω_{γ}^{-1} : submatrix corresponding to $\gamma_k = 1$
- ⊠ Conjugate prior:

$$\beta_{\gamma} | \sigma_{\varepsilon}^2, \gamma \sim N(b_{\gamma}, \sigma_{\varepsilon}^2 \Omega_{\gamma}) \quad \frac{1}{\sigma_{\varepsilon}^2} | \gamma \sim \text{Gamma}\left(\frac{v}{2}, \frac{ss}{2}\right)$$



Full Prior

$$p(\beta, \gamma, \sigma_\varepsilon^2) = p(\beta_\gamma | \gamma, \sigma_\varepsilon^2) p(\sigma_\varepsilon^2 | \gamma) p(\gamma)$$



⊠ ss, v : prior sum of squares and prior sample size



- ⊠ ss, v : prior sum of squares and prior sample size
- ⊠ Expected R^2 , expected model size



Specifying the Prior

- ⊠ ss, v : prior sum of squares and prior sample size
 - ⊠ Expected R^2 , expected model size
- ⊠ π_k : inclusion probability for regression coefficients



Specifying the Prior

- ⊠ ss, v : prior sum of squares and prior sample size
 - ⊠ Expected R^2 , expected model size
- ⊠ π_k : inclusion probability for regression coefficients
- ⊠ Ω^{-1} : based on design matrix X



Specifying the Prior

- ⊠ ss, v : prior sum of squares and prior sample size
 - ⊠ Expected R^2 , expected model size
- ⊠ π_k : inclusion probability for regression coefficients
- ⊠ Ω^{-1} : based on design matrix X
 - ⊠ $\Omega^{-1} = \frac{k}{n} X^T X$



Spike and Slab Posterior

Posterior of β and σ_ε^2 given γ, y^*

⊠ Z_t^* : observation matrix Z_t with $\beta^T x_t = 0$

⊠ $y_t^* = y_t - Z_t^{*T} \alpha_t$

⊠ $y^* = (y_1^*, y_2^*, \dots, y_n^*)^T$

⊠

$$\beta_\gamma | \sigma_\varepsilon, \gamma, y^* \sim N(\tilde{\beta}_\gamma, \sigma_\varepsilon^2 V_\gamma) \quad \frac{1}{\sigma_\varepsilon^2} | \gamma, y^* \sim \text{Gamma}\left(\frac{N}{2}, \frac{SS_\gamma}{2}\right)$$

⊠ $V_\gamma^{-1} = (X^T X)_\gamma + \Omega_\gamma^{-1}$

⊠ $\tilde{\beta}_\gamma = V_\gamma (X_\gamma^T y^* + \Omega_\gamma^{-1} b_\gamma)$

⊠ $N = v + n$

⊠ $SS_\gamma = ss + y^{*T} y^* + b_\gamma^T \Omega_\gamma^{-1} b_\gamma - \tilde{\beta}_\gamma^T V_\gamma^{-1} \tilde{\beta}_\gamma$



Posterior of γ given y^* :

$$\gamma|y^* \sim C(y^*) \frac{|\Omega_\gamma^{-1}|^{1/2}}{|V_\gamma^{-1}|^{1/2}} \frac{p(\gamma)}{SS_\gamma^{N/2-1}}$$



⊠ Combination of:

⊠ stochastic Kalman smoother

⊠ MCMC:

⊠ stochastic search variable selection

⊠ Gibbs sampler