Adventures in Bayesian Structural Time Series

Andrew Bates, Josh Gloyd, Tyler Tucker

What is BSTS?

- Comprised of 3 components:
  - Structural Time Series model (a.k.a. state space model)
  - Spike and Slab regression
  - Bayesian model averaging
- Predicting the Present with Bayesian Structural Time Series by Steven L. Scott and Hal Varian (Google)
- Implementation
  - R: bsts
    - or Causal Impact
  - Python: Causal Impact

Structural Time Series

- Data from unobserved **state space** plus noise
- Model the latent state space instead of the data directly

Local Level Model

- y_t: data
- μ_t: latent state

$$
\begin{aligned}
y_t &= \mu_t + \varepsilon_t & \varepsilon_t &\sim N(0, \sigma_\varepsilon^2) \\
\mu_{t+1} &= \mu_t + \xi_t & \xi_t &\sim N(0, \sigma_\xi^2)
\end{aligned}
$$

Analogous to the inercept in linear regression but allowing for the intercept to vary over time

SAN DIEGO STATE UNIVERSITY

Local Linear Trend Model

- y_t, μ_t: same as before
- ν_t: slope (additional state component)

$$
\begin{aligned}
y_t &= \mu_t + \varepsilon_t & \varepsilon_t &\sim N(0, \sigma_\varepsilon^2) \\
\mu_{t+1} &= \mu_t + \nu_t + \xi_t & \xi_t &\sim N(0, \sigma_\xi^2) \\
\nu_{t+1} &= \nu_t + \zeta_t & \zeta_t &\sim N(0, \sigma_\zeta^2)
\end{aligned}
$$

Structural Time Series

SAN DIEGO STATE UNIVERSITY

General Form

- y_t: data
- α_t: state component

$$
\begin{aligned}
y_t &= Z_t'\alpha_t + \varepsilon_t & \varepsilon_t &\sim N(0, H_t) & (1)\\
\alpha_{t+1} &= T_t\alpha_t + R_t\eta_t & \eta_t &\sim N(0, Q_t) & (2)
\end{aligned}
$$

- (1): **observation equation**
- (2): **transition equation**

Structural Time Series in Bayesian Context

- Spike and slab regression
  - Used when regression components are included
  - Variable selection technique
  - Prior on regression coefficients
- Bayesian Model Averaging
  - Consequence of spike and slab prior
  - Different βs included in each draw of posterior (i.e. different model on each draw)
- Prior Elicitation and Posterior Sampling
  - Inclusion probabilities for regression coefficients
  - Or: expected model size, expected R^2, weight given to R^2
  - Gibbs sampler (stochastic search variable selection) to draw from posterior
  - For details see paper by Scott and Varian