



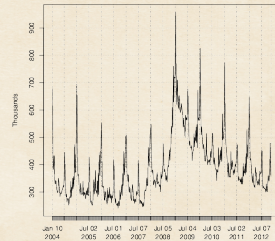
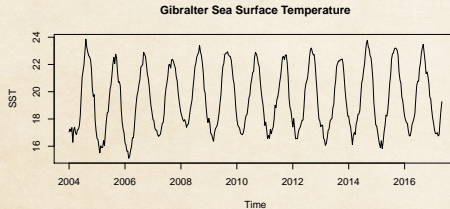
Adventures in Bayesian Structural Time Series

Andrew Bates, Josh Gloyd, Tyler Tucker

What are Time Series?



SAN DIEGO STATE
UNIVERSITY





What are BSTS?

- ⊠ Predicting the Present with Bayesian Structural Time Series
by Steven L. Scott and Hal Varian (Google)



What are BSTS?

- ⊠ Predicting the Present with Bayesian Structural Time Series
by Steven L. Scott and Hal Varian (Google)
- ⊠ Implementation



What are BSTS?

- ⊠ Predicting the Present with Bayesian Structural Time Series
by Steven L. Scott and Hal Varian (Google)
- ⊠ Implementation
 - ⊠ R: [bsts](#)



What are BSTS?

- ⊠ Predicting the Present with Bayesian Structural Time Series
by Steven L. Scott and Hal Varian (Google)
- ⊠ Implementation
 - ⊠ R: [bsts](#)
 - ⊠ or [Causal Impact](#)



What are BSTS?

- ⊠ 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](#)



What are BSTS?

- ⊠ 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](#)
- ⊠ Comprised of 3 components:



What are BSTS?

- ⊗ 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](#)
- ⊗ Comprised of 3 components:
 - ⊗ Structural Time Series model (a.k.a. state space model)



What are BSTS?

- ⊗ 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](#)
- ⊗ Comprised of 3 components:
 - ⊗ Structural Time Series model (a.k.a. state space model)
 - ⊗ Spike and Slab regression



What are BSTS?

- ⊠ 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](#)
- ⊠ Comprised of 3 components:
 - ⊠ Structural Time Series model (a.k.a. state space model)
 - ⊠ Spike and Slab regression
 - ⊠ Bayesian model averaging



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

$$y_t = \mu_t + \varepsilon_t \qquad \varepsilon_t \sim N(0, \sigma_\varepsilon^2)$$

$$\mu_{t+1} = \mu_t + \xi_t \qquad \xi_t \sim N(0, \sigma_\xi^2)$$

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



Local Linear Trend Model

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

$$y_t = \mu_t + \varepsilon_t$$

$$\varepsilon_t \sim N(0, \sigma_\varepsilon^2)$$

$$\mu_{t+1} = \mu_t + \nu_t + \xi_t$$

$$\xi_t \sim N(0, \sigma_\xi^2)$$

$$\nu_{t+1} = \nu_t + \zeta_t$$

$$\zeta_t \sim N(0, \sigma_\zeta^2)$$



General Form

⊠ y_t : data

⊠ α_t : state component

$$y_t = Z_t' \alpha_t + \varepsilon_t \quad \varepsilon_t \sim N(0, H_t) \quad (1)$$

$$\alpha_{t+1} = T_t \alpha_t + R_t \eta_t \quad \eta_t \sim N(0, Q_t) \quad (2)$$

⊠ (1): observation equation

⊠ (2): transition equation



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 β s 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



Joint Prior

$$\otimes \gamma_k = 1 \text{ if } \beta_k \neq 0$$

$$\otimes \gamma_k = 0 \text{ if } \beta_k = 0$$

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



“Spike” Prior

⊠ π_k : probability a particular β_k is included in the model

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



“Slab” Prior

- ⊠ b_γ : b is a vector of prior means
- ⊠ Ω_γ^{-1} : denotes the rows and columns of symmetric matrix Ω^{-1} corresponding to $\gamma_k = 1$
- ⊠ ss : prior sum of squares
- ⊠ ν : prior sample size

$$\beta_\gamma | \sigma_\epsilon^2, \gamma \sim N(b_\gamma, \sigma_\epsilon^2 (\Omega_\gamma^{-1})^{-1}) \quad \frac{1}{\sigma_\epsilon^2} | \gamma \sim Ga\left(\frac{\nu}{2}, \frac{ss}{2}\right)$$