BAYESIAN ECONOMETRICS — HEDIBERT FREITAS LOPES

FINAL EXAM - TAKE HOME — APRIL 11TH 2014

SOLUTION

Regression model with an unknown changepoint. Suppose that the density for y_t, $t = 1, 2, \ldots, T$, can be expressed as

$$(y_t|x_t, \gamma) \sim \begin{cases} N(\theta_1 + \theta_2 x_t, \sigma^2) & t \leq \lambda \\ N(\beta_1 + \beta_2 x_t, \tau^2) & t > \lambda \end{cases},$$

where $\gamma = (\theta_1, \theta_2, \beta_1, \beta_2, \sigma^2, \tau^2, \lambda)'$ and that you employ priors of the form $(\theta_1, \theta_2)' \sim N(\mu_\theta, V_\theta)$, $(\beta_1, \beta_2)' \sim N(\mu_\beta, V_\beta)$, $\sigma^2 \sim IG(a_1, b_1)$, $\tau^2 \sim IG(a_2, b_2)$, and $\lambda \sim U\{1, 2, \ldots, T-1\}$.

Let X be the $n \times 2$ matrix with rows $(1, x_t)$ and $y = (y_1, \ldots, y_n)'$. Also, let $X_{a:b}$ the matrix with rows a to b from X, similarly for $y_{a:b}$. Finally, let $\theta = (\theta_1, \theta_2)'$ and $\beta = (\beta_1, \beta_2)'$. The Gibbs sampler scans through the following 5 steps.

STEP 1: Sample θ from $N(\tilde{\mu}_\theta, \widetilde{V}_\theta)$, where

$$\widetilde{V}_\theta^{-1} = V_\theta^{-1} + X'_{1:\lambda} X_{1:\lambda}/\sigma^2 \quad \text{and} \quad \widetilde{V}_\theta^{-1}\tilde{\mu}_\theta = V_\theta^{-1}\mu_\theta + X'_{1:\lambda} y_{1:\lambda}/\sigma^2.$$

STEP 2: Sample σ^2 from $IG(\tilde{a}_1, \tilde{b}_1)$, where

$$\tilde{a}_1 = a_1 + \tau/2 \quad \text{and} \quad \tilde{b}_1 = b_1 + (y_{1:\lambda} - X_{1:\lambda}\theta)'(y_{1:\lambda} - X_{1:\lambda}\theta)/2.$$

STEP 3: Sample β from $N(\tilde{\mu}_\beta, \widetilde{V}_\beta)$, where

$$\widetilde{V}_\beta^{-1} = V_\beta^{-1} + X'_{\lambda+1:n} X_{\lambda+1:n}/\tau^2 \quad \text{and} \quad \widetilde{V}_\beta^{-1}\tilde{\mu}_\beta = V_\beta^{-1}\mu_\beta + X'_{\lambda+1:n} y_{\lambda+1:n}/\tau^2.$$

STEP 4: Sample σ^2 from $IG(\tilde{a}_1, \tilde{b}_1)$, where

$$\tilde{a}_2 = a_2 + (n-\tau)/2 \quad \text{and} \quad \tilde{b}_2 = b_2 + (y_{\lambda+1:n} - X_{\lambda+1:n}\beta)'(y_{\lambda+1:n} - X_{\lambda+1:n}\beta)/2.$$

STEP 5: Sample λ from $\{1, \ldots, T-1\}$ with probabilities

$$Pr(\lambda|\beta_1, \beta_2, \sigma_1^2, \sigma_2^2, y, x) \propto \prod_{t=1}^{\lambda} p_N(y_t; \beta_1 x_t, \sigma_1^2) \prod_{t=\lambda+1}^{T} p_N(y_t; \beta_2 x_t, \sigma_2^2).$$

Simulation. Generate $T = 500$ observations from the changepoint model $y_t|x_t \sim N(2 + x_t, 0.2)$ for $t \leq 85$ and $N(1.5 + 0.8x_t, 0.5)$, for $t > 85$, where $x_t \sim N(0, 1)$. Select prior hyperparameters as follows: $\mu_\theta = \mu_\beta = 0_2$, $V_\theta = V_\beta = 100I_2$, $a_1 = b_1 = 3$, and $a_2 = b_2 = 0.5$. sample from the joint posterior, $p(\theta_1, \theta_2, \beta_1, \beta_2, \sigma^2, \tau^2, \lambda|y, x)$, via a Gibbs sampler.

Solution: See file `finalexam-simulation-R.pdf`.

Real data. The data below represents yearly homicide rates, y_t, of young men between 15 and 29 years of age (in 100,000 inhabitants) in the city of São Paulo from 1980 to 2009 ($T = 30$).

```
67.01 85.89 82.77 119.73 160.40 161.66 166.90 177.10 161.52 197.68
205.04 210.76 195.83 180.30 198.05 223.14 223.86 220.06 238.41 273.12
267.07 257.76 214.29 227.34 168.58 110.06 83.79 63.47 51.43 56.26
```

MLE of the two regressions, when $\hat{\lambda} = 1998$ is assumed to be the change point, are given by $y_t = 90 + 8x_t$ ($\hat{\sigma} = 20$), for $t \leq 1998$ and $y_t = 822 - 26x_t$ ($\hat{\tau} = 21$), for $t > 1998$, where $x_t = t - 1979$. Assuming the same priors as above, compute the posterior means, standard deviations, medians and 95% posterior credibility intervals of θ_1, θ_2, σ, β_1, β_2, τ and λ.

Solution: See table below and file `finalexam-realdata-R.pdf`.

γ	Mean	Standard Deviation	Median	95% Credibility Interval
θ_1	88.7	8.7	88.7	[72.3;105.7]
θ_2	8.4	0.7	8.4	[7.0;9.7]
σ	18.6	2.7	18.3	[14.3;24.5]
β_1	847.6	92.6	849.3	[644.4;1031.6]
β_2	-27.4	3.5	-27.4	[-34.2;-19.8]
τ	24.5	7.2	23.3	[14.6;41.6]
λ	2000	1.1	2000	[1998;2001]