Regression with Slice Sampling

Andreas Hofheinz

25 11 2020

```
library(bisemSliceSampler)
library(latex2exp)
library(patchwork)
library(ggplot2)
```

```
data(BostonHousing, package  = "mlbench")
keep_cols <- c("crim", "lstat", "age", "medv")
housing_sub <- subset(BostonHousing, select = keep_cols)
```

```
#summary(BostonHousing)
```

```
# Add column with all 1s for intercept
X <- cbind("intercept" = 1, as.matrix(subset(housing_sub, select = -medv)))
y <- housing_sub$medv
```

Model:

$y_{\text{medv },i} = \beta_0 + \beta_1 x_{\text{crim},i} + \beta_2 x_{\text{lstat},i} + \beta_3 x_{\text{age },i} + \beta_4 x_{\text{medv },i} + \epsilon_i, \quad i = 1, \cdots, n$
with $\epsilon_i \overset{\text{iid}}{\sim} \text{Normal}\left(0, \sigma^2\right)$

$\implies y_i \mid \boldsymbol{\beta}, \sigma^2 \overset{\text{i.i.d}}{\sim} \text{N}\left(\mathbf{x_i}^T \beta, \sigma^2\right)$

Likelihood:

$L\left(\mu, \sigma^2\right) \propto \left(\sigma^2\right)^{-n/2} \exp\left(-\frac{1}{2\sigma^2} \sum_{i=1}^{n} \left(y_i - \mathbf{x}_i{}^T \beta\right)^2\right)$

Prior

In this example we use the reference prior for multiple linear regression:

$p\left(\boldsymbol{\beta}, \sigma^2\right) \propto \frac{1}{\sigma^2}$

Posterior

$$p\left(\boldsymbol{\beta}, \sigma^2 \mid \boldsymbol{y}\right) \propto \left(\sigma^2\right)^{-n/2} \exp\left(-\frac{1}{2\sigma^2} \sum_{i=1}^{n} \left(y_i - \mathbf{x}_i{}^T \beta\right)^2\right) * \frac{1}{\sigma^2}$$

Log-Posterior

$$log(p\left(\boldsymbol{\beta}, \sigma^2 \mid \boldsymbol{y}\right)) \propto -\frac{n+2}{2} \log\left(\sigma^2\right) - \frac{1}{2\sigma^2} \sum_{i=1}^{n} \left(y_j - \mathbf{x}_i{}^T \beta\right)^2$$

```
logpost <- function(theta) {
  beta <- theta[-length(theta)]
  sigma_sq <- theta[length(theta)]
  n <- length(y)
  -((n + 2)/2 * log(sigma_sq)) - (1/(2 * sigma_sq)) * crossprod(y - (X %*% beta))
}
```

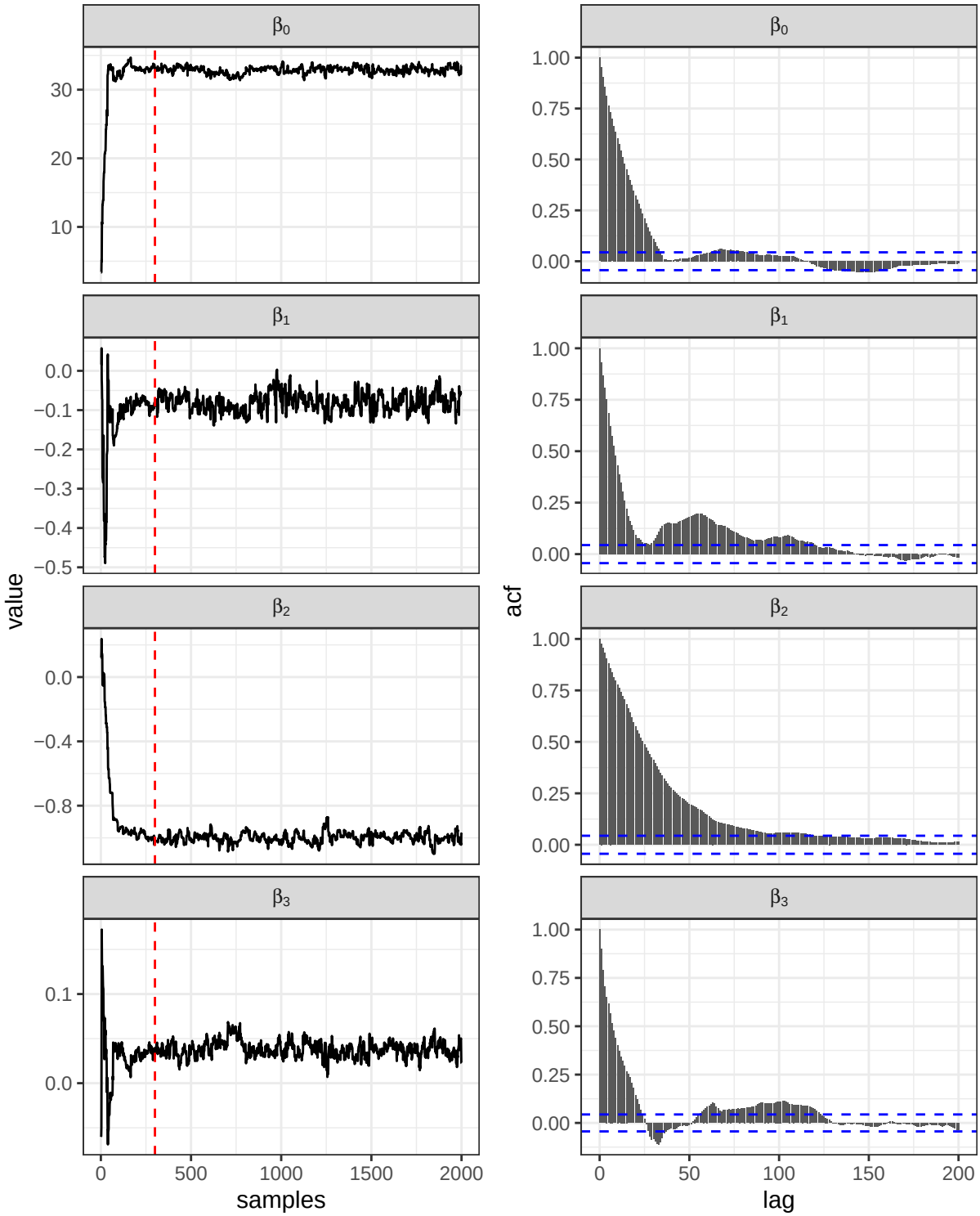
Note: $\boldsymbol{\theta} = (\boldsymbol{\beta}, \sigma^2) = (\beta_0, \beta_1, \beta_2, \beta_3, \sigma^2)$

Sampling from Log-Posterior

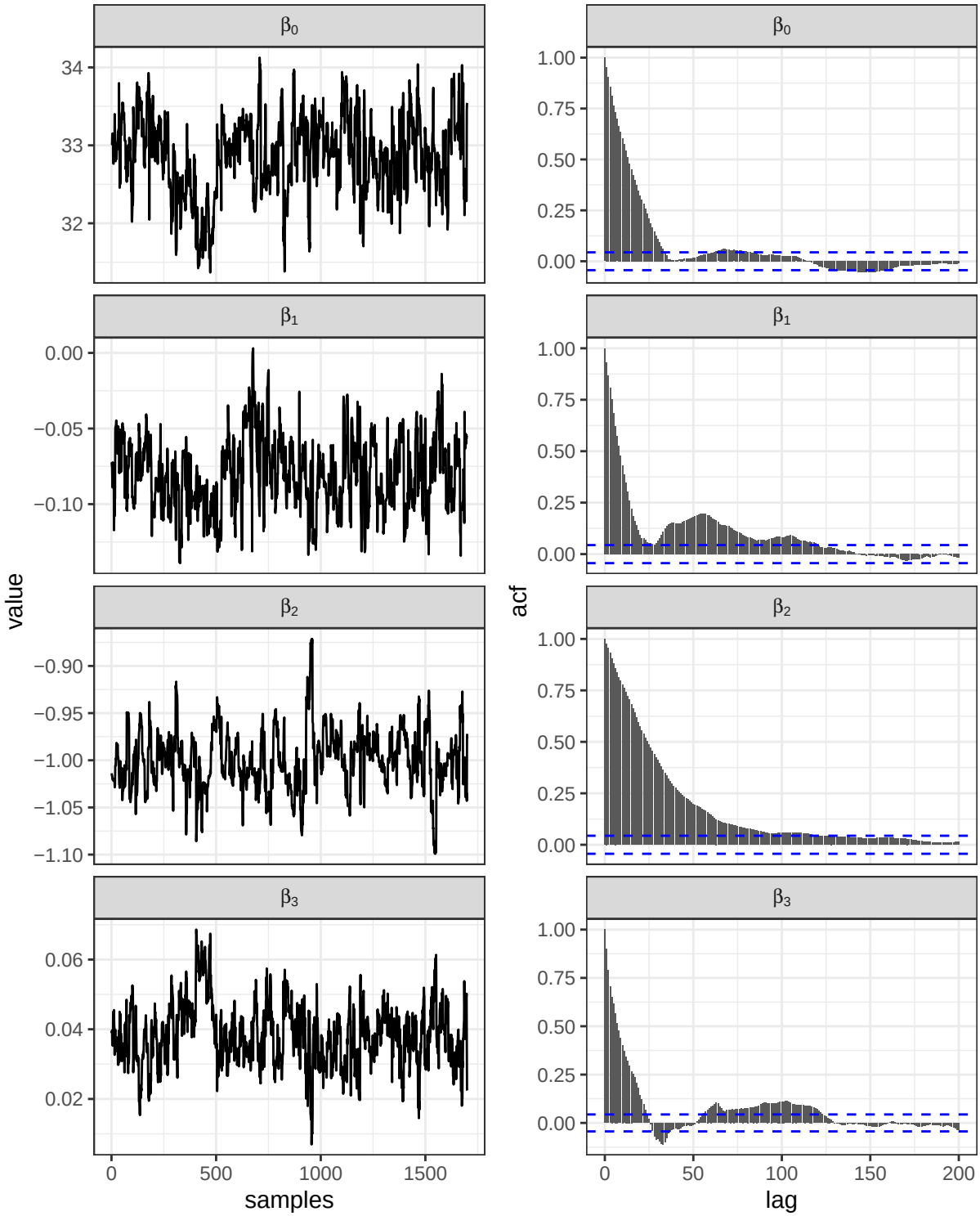
```
w_intercept <- 10
w_rest <- 0.4
w <- c(w_intercept, rep(w_rest, times = ncol(X)))
n_samples <- 2000
theta_init <- c(rep(0, ncol(X)), 10)
```

```
post_samples  <- slice_sampler(
  logpost , x_init = theta_init, w = w, n_samples = n_samples
)
```

Convergence Diagnostic



Convergence Diagnostic - 300 burn-in samples removed



Coefficients and Credibility Intervals

```
            Mean   2.5%  97.5%
intercept 32.858 31.727 33.787
crim      -0.080 -0.124 -0.033
lstat     -1.000 -1.059 -0.939
age        0.038  0.021  0.058
```

Comparison with frequentist calculation

Expected to be similar as non-informative prior was chosen.

```
            Mean   2.5%  97.5%
intercept 32.828 31.359 34.297
crim      -0.083 -0.153 -0.012
lstat     -0.994 -1.094 -0.894
age        0.038  0.014  0.062
```