Homework 5

Data

The data in '`oxboys.csv`' represent the height $y_{i,j}$ of a sample of boys consisting of individuals $i = 1, \ldots, n = 26$ (i.e., subjects) over a period of $J = 9$ ages (centered and scaled). For this homework assignment, we seek to construct a hierarchical model to learn about population-level growth (in height) of boys while considering individual-level heterogeneity. Thus, we only have a single covariate ($x_{i,j} = \texttt{age}_{i,j}$) to use in our Bayesian hierarchical Gaussian model.

Questions

Prepare a written response to the following, using Overleaf. The assignment shouldn't be longer than 10 (double-spaced, excluding title page, references, and appendices). Due Tues., April 11, at the beginning of the class period. Please submit the assignment as a PDF through CANVAS.

1. Using JAGS, develop the R code to fit the Bayesian hierarchical model:

$$y_{i,j} \sim \text{N}(\mathbf{x}'_{i,j}\boldsymbol{\beta}_i, \sigma^2), \text{ for } i = 1, \ldots, n \text{ and } j = 1, \ldots, J \,, \tag{1}$$

$$\boldsymbol{\beta}_i \sim \text{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \,, \tag{2}$$

$$\boldsymbol{\mu} \sim \text{N}(\boldsymbol{\mu}_\mu, \boldsymbol{\Sigma}_\mu) \,, \tag{3}$$

$$(\sigma_0^2)^{-1} \sim \text{Gamma}(a_0, b_0) \,, \tag{4}$$

$$(\sigma_1^2)^{-1} \sim \text{Gamma}(a_1, b_1) \,, \tag{5}$$

$$(\sigma^2)^{-1} \sim \text{Gamma}(a, b) \,, \tag{6}$$

where, $\mathbf{x}_{i,j} = (1, x_{i,j})'$ and

$$\boldsymbol{\beta}_i = \begin{pmatrix} \beta_{0,i} \\ \beta_{1,i} \end{pmatrix} \,, \quad \boldsymbol{\mu} = \begin{pmatrix} \mu_0 \\ \mu_1 \end{pmatrix} \,, \quad \boldsymbol{\Sigma} = \begin{pmatrix} \sigma_0^2 & 0 \\ 0 & \sigma_1^2 \end{pmatrix} \,.$$

2. Use JAGS to fit the Bayesian hierarchical model above to the `oxboys.csv` data set using height as the response variable and age as the predictor variable. Provide 'ridgeline' plots, made using R, to display the marginal posterior distributions of $\boldsymbol{\beta}_i$ and $\boldsymbol{\mu}$ (see Figure 1 below for examples).

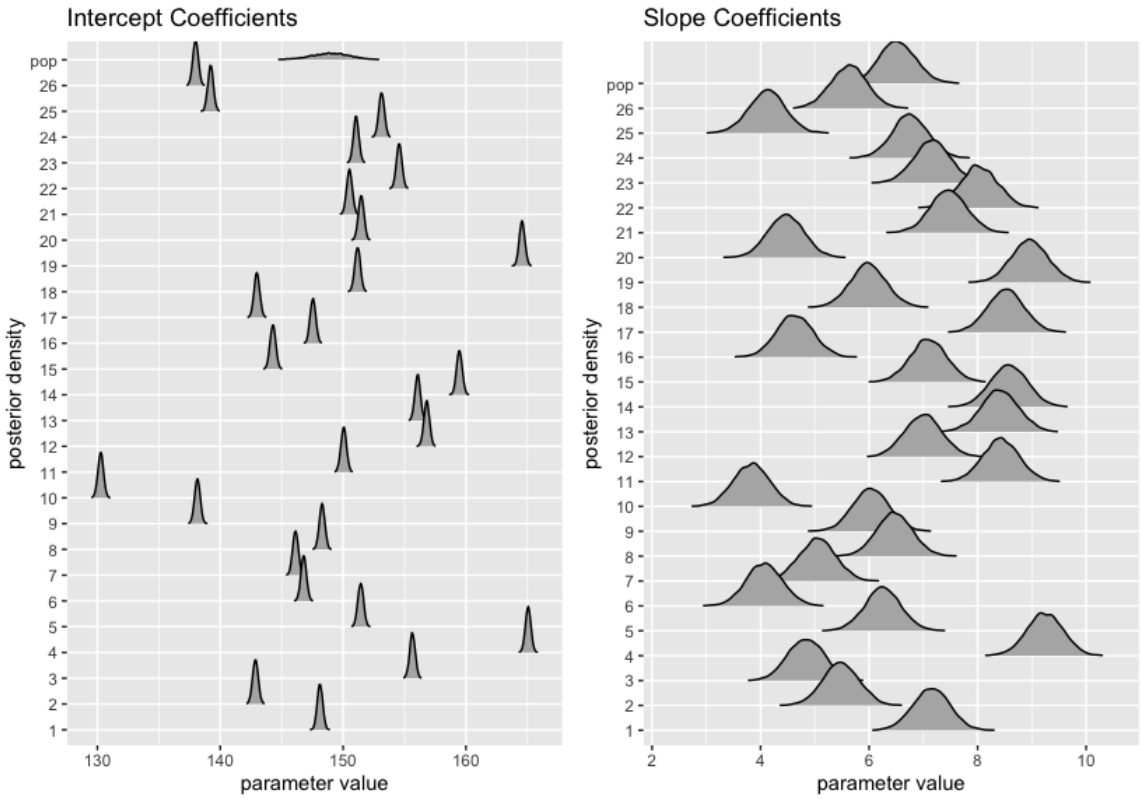

Figure 1: Example ridgeline plots. y-axis labels represent individual (and population-level at top) parameters.

3. Use JAGS to calculate the 'potential scale reduction factors' for all model parameters (including the 'process' variables). You will need to obtain multiple MCMC chains to do this. Report the resulting $\hat{R}$ statistics in a table.

References

- `https://stat.ethz.ch/R-manual/R-devel/library/nlme/html/Oxboys.html`