Bayesian Lasso

John Ormerod

$1 \times p$ $\quad$ $p \times 1$ $\quad$ $i = 1, \ldots, 100.$

1 Problems Week 2

1. Simulate a dataset according to the model

$$y_i = \boldsymbol{\beta}^T \mathbf{x}_i + \varepsilon_i,$$

   with, say, $\mathbf{x}_i \in \mathbb{R}^p$ with $p = 5$ being the number of covariates, $n = 100$ samples, the vector $\boldsymbol{\beta} \in \mathbb{R}^p$ being a sparse vector, and $\varepsilon_i \overset{iid}{\sim} N(0, \sigma^2)$ for some known σ^2.

2. Fit the dataset using some R package that implements lasso regression. Choose a value of λ (the penalty parameter) so that some of the regression coefficients are fitted to be zero.

3. Calculate the full conditional distributions for the model

$$y_i|\boldsymbol{\beta}, \sigma^2 \overset{\text{ind}}{\sim} N(\mathbf{x}_i^T \boldsymbol{\beta}, \sigma^2), \qquad \beta_j|\sigma^2, a_j \sim N\left(0, \frac{\sigma^2 a_j}{\lambda^2}\right), \qquad a_j \overset{\text{iid}}{\sim} \text{Gamma}(1, 1/2), \quad 1 \le j \le p,$$

   with priors $\sigma^2 \sim \text{IG}(a, b)$ and $\lambda > 0$ being the fixed constant chosen in 2, i.e., $p(\boldsymbol{\beta} \,|\, \mathcal{D}, \sigma^2, \mathbf{a})$, $p(\sigma^2 \,|\, \mathcal{D}, \boldsymbol{\beta}, \mathbf{a})$, and $p(a_j \,|\, \mathcal{D}, \boldsymbol{\beta}, \sigma^2)$.

   $\lambda = 2$ $\quad$ x, y $\quad$ x, y

4. Implement a Gibbs sampler Set $\mathbf{a}^{(1)} = \mathbf{1}_n$, $\sigma^{2(1)} = 1$, $\lambda^2 > 0$, $A > 0$, $B > 0$, and $t = 1$. Iterate:

   - $\boldsymbol{\beta}^{(t+1)} \sim p(\boldsymbol{\beta} \mid \mathcal{D}, \sigma^{2(t)}, \mathbf{a}^{(t)})$
   - $\sigma^{2(t+1)} \sim p(\sigma^2 \mid \mathcal{D}, \boldsymbol{\beta}^{(t+1)}, \mathbf{a}^{(t)})$
   - For $j = 1, \ldots, n$ p

$$a_j^{(t+1)} \sim p(a_j \,|\, \mathcal{D}, \boldsymbol{\beta}^{(t+1)}, \sigma^{2(t+1)})$$

   - $t \leftarrow t + 1$.

   Run the sampler until $t = 10^5$ recording all values of $\boldsymbol{\beta}^{(t)}$, $\mathbf{a}^{(t)}$ an $\sigma^{2(t)}$.

5. Plot density plots for each β_j and compare with the fitted values in 2.

2 Problems Week 7

1. The Bayesian EM algorithm iteratively calculates

$$\boldsymbol{\theta}_1^{(t+1)} = \arg\max_{\boldsymbol{\theta}_1} \left[\mathbb{E}_{\boldsymbol{\theta}_2 \,|\, \mathbf{y}, \boldsymbol{\theta}_1^{(t)}} \left[\log p(\mathbf{y}, \boldsymbol{\theta}_1, \boldsymbol{\theta}_2) \right] \right]$$

   Apply the EM algorithm to the Bayesian Lasso model with $\boldsymbol{\theta}_1 = (\boldsymbol{\beta}, \sigma^2)$ and $\boldsymbol{\theta}_2 = \mathbf{a}$.

VB $\quad \hat{A}, \hat{\sigma}^2$