

Linear Regression Model

John Tipton

February 17, 2014

1 Model Statement

1.1 Data Model

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$$

1.2 Process Model

$$\begin{aligned} \boldsymbol{\beta}|\sigma^2, \frac{1}{\gamma^2} &\sim N(0, \Sigma_{\boldsymbol{\beta}}) & \Sigma_{\boldsymbol{\beta}} &= \sigma^2 \mathbf{D}_{\gamma} & \mathbf{D}_{\gamma} &= \text{diag}(\gamma_1, \dots, \gamma_p) \\ \boldsymbol{\epsilon} &\sim N(0, \Sigma_{\boldsymbol{\epsilon}}) & \Sigma_{\boldsymbol{\epsilon}} &= \sigma_{\epsilon}^2 \mathbf{I} \end{aligned}$$

1.3 Parameter Model

$$\begin{aligned} \sigma_{\epsilon}^2 &\sim \text{IG}(\alpha_{\epsilon}, \beta_{\epsilon}) \\ \gamma_j &\sim \text{Exponential}(\lambda^2/2) \text{ for each } j = 1, \dots, p \\ \lambda^2 &\sim \text{Gamma}(\alpha_{\lambda}, \beta_{\lambda}) \end{aligned}$$

where p is the number of parameters in $\boldsymbol{\beta}$ and $\boldsymbol{\mu}_{\boldsymbol{\beta}} = 0$. Equivalently, $1/\gamma_j^2 \sim \text{IG}(1, \lambda^2/2)$.

2 Posterior

$$[\boldsymbol{\beta}, \gamma^2, \sigma_{\epsilon}^2, \lambda^2 | \mathbf{y}] \propto [\mathbf{y} | \boldsymbol{\beta}, \sigma_{\epsilon}^2][\boldsymbol{\beta} | \sigma_{\epsilon}^2, \gamma^2][\sigma_{\epsilon}^2][\gamma][\lambda^2]$$

3 Full Conditionals

3.1 Full Conditional for $\boldsymbol{\beta}$

$$\begin{aligned} [\boldsymbol{\beta} | \cdot] &\propto [\mathbf{y} | \boldsymbol{\beta}, \sigma_{\epsilon}^2][\boldsymbol{\beta} | \sigma_{\epsilon}^2, \gamma^2] \\ &\propto e^{-\frac{1}{2}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})^T \Sigma_{\epsilon}^{-1}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})} e^{-\frac{1}{2}\boldsymbol{\beta}^T \Sigma_{\boldsymbol{\beta}}^{-1} \boldsymbol{\beta}} \\ &\propto e^{-\frac{1}{2}(\boldsymbol{\beta}^T (\mathbf{X}^T \Sigma_{\epsilon}^{-1} \mathbf{X} + \Sigma_{\boldsymbol{\beta}}^{-1}) \boldsymbol{\beta} - 2\boldsymbol{\beta}^T (\mathbf{X}^T \Sigma_{\epsilon}^{-1} \mathbf{y} + \Sigma_{\boldsymbol{\beta}}^{-1} \boldsymbol{\mu}_{\boldsymbol{\beta}}))} \end{aligned}$$

which is Normal with mean $\mathbf{A}^{-1}\mathbf{b}$ and variance \mathbf{A}^{-1} where

$$\begin{aligned}\mathbf{A}^{-1} &= (\mathbf{X}^T \boldsymbol{\Sigma}_\epsilon^{-1} \mathbf{X} + \boldsymbol{\Sigma}_\beta^{-1})^{-1} \\ &= (\mathbf{X}^T \mathbf{X} + \mathbf{D}_\gamma)^{-1} \\ \mathbf{b} &= (\mathbf{X}^T \boldsymbol{\Sigma}_\epsilon^{-1} \mathbf{y} + \boldsymbol{\Sigma}_\beta^{-1} \boldsymbol{\mu}_\beta) \\ &= \mathbf{X}^T \mathbf{y}\end{aligned}$$

3.2 Full Conditional for σ_ϵ^2

$$\begin{aligned}[\sigma_\epsilon^2 | \cdot] &\propto [\mathbf{y} | \boldsymbol{\beta}, \sigma_\epsilon^2][\boldsymbol{\beta} | \sigma_\epsilon^2, \gamma^2][\sigma_\epsilon^2] \\ &\propto (|\boldsymbol{\Sigma}_\epsilon|^{-\frac{1}{2}}) e^{-\frac{1}{2}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})^T \boldsymbol{\Sigma}_\epsilon^{-1}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})} (|\boldsymbol{\Sigma}_\beta|^{-\frac{1}{2}}) e^{-\frac{1}{2}\boldsymbol{\beta}^T \boldsymbol{\Sigma}_\beta^{-1} \boldsymbol{\beta}} (\sigma_\epsilon^2)^{-\alpha_\epsilon+1} e^{-\frac{\beta_\epsilon}{\sigma_\epsilon^2}} \\ &\propto (\sigma_\epsilon^2)^{-\frac{n}{2} - \frac{p}{2} - \alpha_\epsilon - 1} e^{-\frac{1}{\sigma_\epsilon^2}(\frac{1}{2}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})^T(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}) + \frac{1}{2}\boldsymbol{\beta}^T \mathbf{D}_\gamma^{-1} \boldsymbol{\beta} + \beta_\epsilon)}\end{aligned}$$

IG($\alpha_\beta + \frac{n}{2} + \frac{p}{2}, \frac{1}{2}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})^T(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}) + \frac{1}{2}\boldsymbol{\beta}^T \mathbf{D}_\gamma \boldsymbol{\beta} + \beta_\beta$) since $|\boldsymbol{\Sigma}_\epsilon| = (\sigma_\epsilon^2)^n$, $\boldsymbol{\Sigma}_\epsilon^{-1} = \frac{1}{\sigma_\epsilon^2} \mathbf{I}$, $|\boldsymbol{\Sigma}_\beta| = (\sigma_\epsilon^2)^p |\mathbf{D}_\gamma|$, and $\boldsymbol{\Sigma}_\beta^{-1} = \frac{1}{\sigma_\epsilon^2} \mathbf{D}_\gamma$

3.3 Full Conditional for $\frac{1}{\gamma_j}$

$$\begin{aligned}[\frac{1}{\gamma_j^2} | \cdot] &\propto [\beta_j | \sigma_\epsilon^2, \gamma_j^2][\frac{1}{\gamma_j^2}] \\ &\propto |\gamma_j^2|^{-\frac{1}{2}} e^{-\frac{1}{\gamma_j^2} \frac{\beta_j^2}{2\sigma_\epsilon^2}} (\frac{1}{\gamma_j^2})^{-2} e^{\frac{\lambda^2/2}{\gamma_j^2}} \\ &\propto (\frac{1}{\gamma_j^2})^{-3/2} e^{-\frac{1}{2}(\frac{1}{\gamma_j^2} \frac{\beta_j^2}{\sigma_\epsilon^2} + \frac{\lambda^2}{\gamma_j^2})}\end{aligned}$$

which is Inverse Gaussian(μ', λ') where $\mu' = \sqrt{\frac{\lambda^2 \sigma_\epsilon^2}{\beta_j^2}}$ and $\lambda' = \lambda^2$. The inverse Gaussian distribution for $x > 0, \mu' > 0$, and $\lambda > 0$

$$f(x) = (\frac{\lambda^2}{2\pi x^3})^{1/2} e^{-\frac{\lambda'(x-\mu')^2}{2\mu'^2 x}}$$

3.4 Full Conditional for λ^2

$$\begin{aligned}[\lambda^2 | \cdot] &\propto [\gamma^2 | \lambda^2][\lambda^2] \\ &\propto \prod_{j=1}^p \frac{\lambda^2}{2} e^{-\lambda^2 \frac{\gamma_j^2}{2}} (\lambda^2)^{\alpha_\lambda - 1} e^{-\beta_\lambda \lambda^2} \\ &\propto (\lambda^2)^{\alpha_\lambda + p + 1} e^{-\lambda^2(\beta_\lambda + \sum_{j=1}^p \gamma_j^2)}\end{aligned}$$

which is Gamma($\alpha_\lambda + p, \beta_\lambda + \frac{1}{2} \sum_{j=1}^p \gamma_j^2$)

4 Posterior Predictive Distribution

The posterior predictive distribution for \mathbf{y}_t is sampled at each MCMC iteration k by

$$\mathbf{y}_t^{(k)} \sim N(\mathbf{H}_t \mathbf{X} \boldsymbol{\beta}_t^{(k)}, \boldsymbol{\Sigma}^{(k)})$$