Linear Regression Model

John Tipton

February 5, 2014

1 Model Statement

1.1 Data Model

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

1.2 Process Model

$$\begin{aligned} \boldsymbol{\beta}|\sigma^2 &\sim N(0, \sigma^2 \boldsymbol{D}_\gamma) & \boldsymbol{D}_\gamma &= diag(\gamma_1, \ldots, \gamma_p) \\ \boldsymbol{\epsilon} &\sim N(0, \boldsymbol{\Sigma}_\epsilon) & \boldsymbol{\Sigma}_\epsilon &= \sigma_\epsilon^2 \boldsymbol{I} \end{aligned}$$

1.3 Parameter Model

$$\begin{aligned} &\sigma_\epsilon^2 \sim IG(\alpha_\epsilon, \beta_\epsilon) \\ &\gamma_j \sim Exp(\lambda^2/2) foreach j = 1, \ldots, p \\ &\lambda^2 \sim Gamma(\alpha_\lambda, \beta_\lambda) \end{aligned}$$

where $\boldsymbol{I}_\beta$ is the identity matrix of size $\tau \times \tau$ where τ is the number of parameters in $\boldsymbol{\beta}$, $\boldsymbol{I}$ is the identity matrix of size $n \times n$ and n is the number of samples of y

2 Posterior

$$[\boldsymbol{\beta}, \boldsymbol{\mu}_\beta, \sigma_\beta^2, \sigma_\epsilon^2 | \boldsymbol{y}] \propto [\boldsymbol{y}|\boldsymbol{\beta}, \sigma_\epsilon^2][\boldsymbol{\beta}|\boldsymbol{\mu}_\beta, \sigma_\beta^2][\boldsymbol{\mu}_\beta][\sigma_\beta^2][\sigma_\epsilon^2]$$

3 Full Conditionals

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

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

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

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

3.2 Full Conditional for $\boldsymbol{\mu}_\beta$

$$\begin{aligned}
[\boldsymbol{\mu}_\beta|\cdot] &\propto [\boldsymbol{\beta}|\boldsymbol{\mu}_\beta, \sigma_\beta^2][\boldsymbol{\mu}_\beta] \\
&\propto e^{-\frac{1}{2}(\boldsymbol{\beta}-\boldsymbol{\mu}_\beta)^T\boldsymbol{\Sigma}_\beta^{-1}(\boldsymbol{\beta}-\boldsymbol{\mu}_\beta)}e^{-\frac{1}{2}(\boldsymbol{\mu}_\beta-\boldsymbol{\mu}_0)^T\Sigma_0^{-1}(\boldsymbol{\mu}_\beta-\boldsymbol{\mu}_0)} \\
&\propto e^{-\frac{1}{2}(\boldsymbol{\mu}_\beta^T(\boldsymbol{\Sigma}_\beta^{-1}+\boldsymbol{\Sigma}_0^{-1})\boldsymbol{\mu}_\beta-2\boldsymbol{\mu}_\beta^T(\boldsymbol{\Sigma}_\beta^{-1}\boldsymbol{\beta}+\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\mu}_0))}
\end{aligned}$$

which is multivariate normal with mean $\boldsymbol{A}^{-1}\boldsymbol{b}$ and variance $\boldsymbol{A}^{-1}$ where

$$\begin{aligned}
\boldsymbol{A}^{-1} &= (\boldsymbol{\Sigma}_\beta^{-1} + \boldsymbol{\Sigma}_0^{-1})^{-1} \\
\boldsymbol{b} &= (\boldsymbol{\Sigma}_\beta^{-1}\boldsymbol{\beta} + \boldsymbol{\Sigma}_0^{-1}\boldsymbol{\mu}_0)
\end{aligned}$$

3.3 Full Conditional for σ_β^2

$$\begin{aligned}
[\sigma_\beta^2|\cdot] &\propto [\boldsymbol{\beta}|\boldsymbol{\mu}_\beta, \sigma_\beta^2][\sigma_\beta^2] \\
&\propto |\boldsymbol{\Sigma}_\beta|^{-\frac{1}{2}})e^{-\frac{1}{2}(\boldsymbol{\beta}-\boldsymbol{\mu}_\beta)^T\boldsymbol{\Sigma}_\beta^{-1}(\boldsymbol{\beta}-\boldsymbol{\mu}_\beta)}(\sigma_\beta^2)^{-(\alpha_\beta+1)}e^{-\frac{\beta_\beta}{\sigma_\beta^2}} \\
&\propto (\sigma_\beta^2)^{-(\alpha_\beta+\frac{\tau}{2}+1)}e^{-\frac{1}{\sigma_\beta^2}(\frac{1}{2}(\boldsymbol{\beta}-\boldsymbol{\mu}_\beta)^T(\boldsymbol{\beta}-\boldsymbol{\mu}_\beta)+\beta_\beta)}
\end{aligned}$$

which is $\text{IG}(\alpha_\beta + \frac{\tau}{2}, \frac{1}{2}(\boldsymbol{\beta} - \boldsymbol{\mu}_\beta)^T(\boldsymbol{\beta} - \boldsymbol{\mu}_\beta) + \beta_\beta)$ since the determinant $|\boldsymbol{\Sigma}_\beta| = (\sigma_\beta^2)^\tau$ and $\boldsymbol{\Sigma}_\beta^{-1} = \frac{1}{\sigma_\beta^2}\boldsymbol{I}$

3.4 Full Conditional for σ_ϵ^2

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

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

4 Posterior Predictive Distribution

The posterior predictive distribution for $\boldsymbol{y}_t$ is sampled a each MCMC iteration k by

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