

Linear Regression Model

John Tipton

January 25, 2014

1 Model Statement

1.1 Data Model

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

1.2 Process Model

$$\begin{aligned}\boldsymbol{\beta} &\sim N(\boldsymbol{\mu}_\beta, \boldsymbol{\Sigma}_\beta) & \boldsymbol{\Sigma}_\beta &= \sigma_\beta^2 \mathbf{I} \\ \boldsymbol{\epsilon} &\sim N(0, \boldsymbol{\Sigma}_\epsilon) & \boldsymbol{\Sigma}_\epsilon &= \sigma_\epsilon^2 \mathbf{I}\end{aligned}$$

1.3 Parameter Model

$$\begin{aligned}\boldsymbol{\mu}_\beta &\sim N(\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0) \\ \sigma_\beta^2 &\sim IG(\alpha_\beta, \beta_\beta) \\ \sigma_\epsilon^2 &\sim IG(\alpha_\epsilon, \beta_\epsilon)\end{aligned}$$

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

2 Posterior

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

3 Full Conditionals

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

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

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

$$\begin{aligned}\mathbf{A}^{-1} &= (\mathbf{X}^T \Sigma_\epsilon^{-1} \mathbf{X} + \Sigma_\beta^{-1})^{-1} \\ \mathbf{b} &= (\mathbf{X}^T \Sigma_\epsilon^{-1} \mathbf{y} + \Sigma_\beta^{-1} \mu_\beta)\end{aligned}$$

3.2 Full Conditional for μ_β

$$\begin{aligned}[\mu_\beta | \cdot] &\propto [\beta | \mu_\beta, \sigma_\beta^2][\mu_\beta] \\ &\propto e^{-\frac{1}{2}(\beta - \mu_\beta)^T \Sigma_\beta^{-1} (\beta - \mu_\beta)} e^{-\frac{1}{2}(\mu_\beta - \mu_0)^T \Sigma_0^{-1} (\mu_\beta - \mu_0)} \\ &\propto e^{-\frac{1}{2}(\mu_\beta^T (\Sigma_\beta^{-1} + \Sigma_0^{-1}) \mu_\beta - 2\mu_\beta^T (\Sigma_\beta^{-1} \beta + \Sigma_0^{-1} \mu_0))}\end{aligned}$$

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

$$\begin{aligned}\mathbf{A}^{-1} &= (\Sigma_\beta^{-1} + \Sigma_0^{-1})^{-1} \\ \mathbf{b} &= (\Sigma_\beta^{-1} \beta + \Sigma_0^{-1} \mu_0)\end{aligned}$$

3.3 Full Conditional for σ_β^2

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

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

3.4 Full Conditional for σ_ϵ^2

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

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

4 Posterior Predictive Distribution

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

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