Linear Regression with Sets of Conjugate Priors

Gero Walter

Department of Statistics
Ludwig-Maximilians-University Munich

September 14th, 2009

Linear Regression

$$z_i = \mathbf{x}_i^\mathsf{T}\beta + \varepsilon_i\,, \quad \mathbf{x}_i \in \mathbb{R}^p\,,\ \beta \in \mathbb{R}^p\,,\ \varepsilon_i \sim \mathrm{N}(0,\sigma^2)\,,$$

$$z = \mathbf{X}\beta + \varepsilon \qquad \rightarrow \qquad z \sim \mathrm{N}(\mathbf{X}\beta, \sigma^2\mathbf{I})$$

Least squares: $\hat\beta_{LS} = (\mathbf{X}^\mathsf{T}\mathbf{X})^{-1}\mathbf{X}^\mathsf{T}z$, with

$$\begin{aligned}
\mathbb{E}[\hat\beta_{LS}] &= \beta\,,\\
\mathbb{V}(\hat\beta_{LS}) &= \hat\sigma^2_{LS}(\mathbf{X}^\mathsf{T}\mathbf{X})^{-1}\,,\\
\hat\sigma^2_{LS} &= \frac{1}{n-p}(z-\mathbf{X}\hat\beta_{LS})^\mathsf{T}(z-\mathbf{X}\hat\beta_{LS})
\end{aligned}$$

Standard Conjugate Prior (SCP)

$$p(\beta, \sigma^2) = p(\beta \mid \sigma^2)p(\sigma^2)$$

$$p(\beta \mid \sigma^2) \sim \mathrm{N}\left(m^{(0)}, \sigma^2 \mathbf{M}^{(0)}\right)$$

$$m^{(0)} \in \mathbb{R}^p, \; \mathbf{M}^{(0)} \in \mathbb{R}^{p \times p} \text{ symm. pos. def.}$$

$$p(\sigma^2) \sim \mathrm{IG}\left(a^{(0)}, b^{(0)}\right) \qquad a^{(0)} > 0, \; b^{(0)} > 0$$

Standard Conjugate Prior (SCP)

Updating:

$$
\begin{aligned}
m^{(1)} &= \left(\mathbf{M}^{(0)^{-1}} + \mathbf{X}^\mathsf{T}\mathbf{X}\right)^{-1} \left(\mathbf{M}^{(0)^{-1}} m^{(0)} + \mathbf{X}^\mathsf{T} z\right) , \\
\mathbf{M}^{(1)} &= \left(\mathbf{M}^{(0)^{-1}} + \mathbf{X}^\mathsf{T}\mathbf{X}\right)^{-1} , \\
a^{(1)} &= a^{(0)} + \frac{n}{2} \\
b^{(1)} &= b^{(0)} + \frac{1}{2}\left(z^\mathsf{T} z + m^{(0)^\mathsf{T}} \mathbf{M}^{(0)^{-1}} m^{(0)} - m^{(1)^\mathsf{T}} \mathbf{M}^{(1)^{-1}} m^{(1)}\right) .
\end{aligned}
$$

$$
\mathbb{E}[\sigma^2 \mid z] = \frac{2a^{(0)} - 2}{2a^{(0)} + n - 2}\mathbb{E}[\sigma^2] + \frac{n-p}{2a^{(0)} + n - 2}\hat{\sigma}^2_{LS} + \frac{p}{2a^{(0)} + n - 2}\hat{\sigma}^2_{PDC}
$$

Canonical Construction of Conjugate Priors

If likelihood has linear, canonical exponential family form

$$f(z \mid \psi) = \mathbf{a}(z) \cdot \exp\{\langle \psi, \tau(z) \rangle - n \cdot \mathbf{b}(\psi)\}\,,$$

➡ a conjugate prior is:

$$p(\psi) = \mathbf{c}(n^{(0)}, y^{(0)}) \cdot \exp\left\{ n^{(0)} \cdot [\langle \psi, y^{(0)} \rangle - \mathbf{b}(\psi)] \right\}$$

➡ posterior: $y^{(0)} \longrightarrow y^{(1)}$, $n^{(0)} \longrightarrow n^{(1)}$, where

$$n^{(1)} = n^{(0)} + n$$

$$y^{(1)} = \frac{n^{(0)} y^{(0)} + \tau(z)}{n^{(0)} + n} = \frac{n^{(0)}}{n^{(0)} + n} y^{(0)} + \frac{n}{n^{(0)} + n} \frac{1}{n} \tau(z)$$

Canonically Constructed Conjugate Prior (CCCP)

Likelihood $z \sim \mathrm{N}(\mathbf{X}\beta, \sigma^2\mathbf{I})$ has indeed linear, canonical exponential family form

$$\psi = \begin{pmatrix} \frac{\beta}{\sigma^2} \\ -\frac{1}{\sigma^2} \end{pmatrix}, \quad \tau(z) = \begin{pmatrix} \mathbf{X}^\mathsf{T} z \\ \frac{1}{2} z^\mathsf{T} z \end{pmatrix}, \quad \mathbf{b}(\psi) = \frac{1}{2n\sigma^2}\beta^\mathsf{T}\mathbf{X}^\mathsf{T}\mathbf{X}\beta + \frac{1}{2}\log(\sigma^2)\,.$$

- ➡ constructed prior is on $\psi = \psi(\beta, \sigma^2)$
- ➡ must be transformed, as we want a prior on $\theta = (\beta,\, \sigma^2)^\mathsf{T}$
- ➡ ... ➡

$$\beta \mid \sigma^2 \sim \mathrm{N}_p\Big(\underbrace{n\,(\mathbf{X}^\mathsf{T}\mathbf{X})^{-1}\,y_1^{(0)}}_{m^{(0)}},\ \sigma^2 \underbrace{\frac{n}{n^{(0)}}(\mathbf{X}^\mathsf{T}\mathbf{X})^{-1}}_{\mathbf{M}^{(0)}}\Big)$$

$$\sigma^2 \sim \mathrm{IG}\Big(\underbrace{\frac{n^{(0)}+p+2}{2}}_{a^{(0)}},\ \underbrace{n^{(0)}y_2^{(0)} - \frac{n^{(0)}}{2}y_1^{(0)\mathsf{T}}\,n(\mathbf{X}^\mathsf{T}\mathbf{X})^{-1}y_1^{(0)}}_{b^{(0)}}\Big)\,.$$

Update of $\beta \mid \sigma^2$

$$
\begin{aligned}
\mathbb{E}[\beta \mid \sigma^2, z] &= m^{(1)} = n(\mathbf{X}^\mathsf{T}\mathbf{X})^{-1}\, y_1^{(1)} \\
&= n(\mathbf{X}^\mathsf{T}\mathbf{X})^{-1}\left(\frac{n^{(0)}}{n^{(0)}+n}\, y_1^{(0)} + \frac{n}{n^{(0)}+n}\cdot\frac{1}{n}(\mathbf{X}^\mathsf{T} z)\right) \\
&= n(\mathbf{X}^\mathsf{T}\mathbf{X})^{-1}\frac{n^{(0)}}{n^{(0)}+n}\cdot\frac{1}{n}(\mathbf{X}^\mathsf{T}\mathbf{X})m^{(0)} \\
&\qquad\qquad + n(\mathbf{X}^\mathsf{T}\mathbf{X})^{-1}\frac{n}{n^{(0)}+n}\cdot\frac{1}{n}(\mathbf{X}^\mathsf{T} z) \\
&= \frac{n^{(0)}}{n^{(0)}+n}\mathbb{E}[\beta \mid \sigma^2] + \frac{n}{n^{(0)}+n}\hat{\beta}_{LS}
\end{aligned}
$$

$$
\mathbb{V}(\beta \mid \sigma^2, z) = \sigma^2\frac{n}{n^{(1)}}(\mathbf{X}^\mathsf{T}\mathbf{X})^{-1} = \sigma^2\frac{n}{n^{(0)}+n}(\mathbf{X}^\mathsf{T}\mathbf{X})^{-1}
$$

Update of σ^2

$$\mathbb{E}[\sigma^2] = \frac{b^{(0)}}{a^{(0)} - 1} = \frac{2n^{(0)}}{n^{(0)} + p} y_2^{(0)} - \frac{1}{n^{(0)} + p} m^{(0)^\mathsf{T}} \mathbf{M}^{(0)^{-1}} m^{(0)} .$$

$$\begin{aligned}\mathbb{E}[\sigma^2 \mid z] = {} & \frac{n^{(0)} + p}{n^{(0)} + n + p} \mathbb{E}[\sigma^2] + \frac{n - 1}{n^{(0)} + n + p} \frac{1}{n - 1} z^\mathsf{T} z \\ & + \frac{1}{n^{(0)} + n + p} \Big(m^{(0)^\mathsf{T}} \mathbf{M}^{(0)^{-1}} m^{(0)} - m^{(1)^\mathsf{T}} \mathbf{M}^{(1)^{-1}} m^{(1)} \Big) ,\end{aligned}$$

Update of σ^2

$$\mathbb{E}[\sigma^2 \mid z] = \frac{n^{(0)} + p}{n^{(0)} + n + p}\mathbb{E}[\sigma^2] + \frac{n - p}{n^{(0)} + n + p}\hat{\sigma}^2_{LS} + \frac{p}{n^{(0)} + n + p}\sigma^2_{PDC}\,,$$

$$\mathbb{E}[\sigma^2 \mid z] = \frac{n^{(0)} + p}{n^{(0)} + n + p}\mathbb{E}[\sigma^2] + \frac{n^{(0)} + p}{n^{(0)} + n + p}{\sigma^{(0)}}^2 + \frac{2(n - p)}{n^{(0)} + n + p}\hat{\sigma}^2_{LS} - \frac{n^{(1)} + p}{n^{(0)} + n + p}{\sigma^{(1)}}^2\,.$$

$$\sigma^2_{PDC} = \frac{n}{n^{(0)} + n}\frac{1}{p}(m^{(0)} - \hat{\beta}_{LS})^\mathsf{T}{\mathbf{M}^{(0)}}^{-1}(m^{(0)} - \hat{\beta}_{LS})\,,$$

$${\sigma^{(0)}}^2 = \frac{n^{(0)}}{n} \cdot \frac{1}{n^{(0)} + p}(z - \mathbf{X}m^{(0)})^\mathsf{T}(z - \mathbf{X}m^{(0)})$$

$${\sigma^{(1)}}^2 = \frac{n^{(1)}}{n} \cdot \frac{1}{n^{(1)} + p}(z - \mathbf{X}m^{(1)})^\mathsf{T}(z - \mathbf{X}m^{(1)})$$

all of which satisfy $\mathbb{E}[\,\cdot \mid \sigma^2] = \sigma^2$.

Update of β

$$\mathbb{E}[\beta \mid z] = \mathbb{E}[\beta \mid \sigma^2, z] = \frac{n^{(0)}}{n^{(0)} + n} \mathbb{E}[\beta \mid \sigma^2] + \frac{n}{n^{(0)} + n} \hat{\beta}_{LS}$$

$$\mathbb{V}(\beta \mid z) = \frac{b^{(1)}}{a^{(1)} - 1} \mathbf{M}^{(1)} = \mathbb{E}[\sigma^2 \mid z] \, \frac{n}{n^{(1)}} (\mathbf{X}^\mathsf{T}\mathbf{X})^{-1}$$

(generalized) iLUCK-model-calculous

iLUCK-model:

$$\left\{\left(n^{(1)}, y^{(1)}\right) \,\middle|\, n^{(1)} = n^{(0)} + n,\ y^{(1)} = \frac{n^{(0)} y^{(0)} + \tau(x)}{n^{(0)} + n},\ y^{(0)} \in \mathcal{Y}^{(0)}\right\}$$

generalized iLUCK-model:

$$\left\{\left(n^{(1)}, y^{(1)}\right) \,\middle|\, n^{(1)} = n^{(0)} + n,\ y^{(1)} = \frac{n^{(0)} y^{(0)} + \tau(x)}{n^{(0)} + n},\ n^{(0)} \in \mathcal{N}^{(0)},\ y^{(0)} \in \mathcal{Y}^{(0)}\right\}$$

First regression lines