Inconsistency of Bayesian Inference for Misspecified Linear Models, and a Proposal for Repairing It — ERRATUM

Peter Grünwald   Thijs van Ommen

October 29, 2018

In the journal paper, at the end of Section 2.5, we give an expression for the parameter $b_{n,\eta}$ of the σ^2-component of the generalized Bayesian posterior. This expression is incorrect. Thanks go to Tom Viering for spotting this error by noticing that the expression did not agree with standard Bayes when $\eta = 1$! The correct expression is

$$b_{n,\eta} = b_0 + \frac{1}{2}\bar{\beta}_0^T \Sigma_0^{-1} \bar{\beta}_0 + \frac{\eta}{2}(y^n)^T y^n - \frac{1}{2}\bar{\beta}_{n,\eta}^T \Sigma_{n,\eta}^{-1} \bar{\beta}_{n,\eta}.$$

Our code calculated $b_{n,\eta}$ in a different manner, so the experimental results remain valid: they are not affected by this mistake.

Derivation Write $\Sigma^{-1} = \mathbf{X}_0^T \mathbf{X}_0$. For a prior π given by $N(\bar{\beta}_0, \sigma^2(\mathbf{X}_0^T\mathbf{X}_0)^{-1})$ and Inv-gamma$(\sigma^2 \mid a_0, b_0)$, the generalized posterior is proportional to

$$\begin{aligned}
\pi(\beta, \sigma \mid y^n, \eta) &\propto (f_{\beta,\sigma^2}(y^n))^\eta \pi(\beta \mid \sigma)\pi(\sigma) \\
&\propto \sigma^{-\eta n} \exp\left[-\frac{\eta}{2\sigma^2}(y^n - \mathbf{X}_n\beta)^T(y^n - \mathbf{X}_n\beta)\right] \\
&\quad \cdot \sigma^{-p} \exp\left[-\frac{1}{2\sigma^2}(\beta - \bar{\beta}_0)^T \mathbf{X}_0^T \mathbf{X}_0 (\beta - \bar{\beta}_0)\right] \\
&\quad \cdot \sigma^{-2(a_0+1)} \exp\left[-\frac{b_0}{\sigma^2}\right].
\end{aligned}$$

Defining

$$\mathbf{X}^* = \begin{bmatrix} \mathbf{X}_0 \\ \sqrt{\eta}\mathbf{X}_n \end{bmatrix}, y^* = \begin{bmatrix} \mathbf{X}_0\bar{\beta}_0 \\ \sqrt{\eta} y^n \end{bmatrix}, \Sigma_{n,\eta} = (\mathbf{X}^{*T}\mathbf{X}^*)^{-1}, \text{ and } \bar{\beta}_{n,\eta} = (\mathbf{X}^{*T}\mathbf{X}^*)^{-1}\mathbf{X}^{*T}y^*,$$

this can be rewritten as

$$\begin{aligned}
\pi(\beta, \sigma \mid y^n, \eta) &\propto \sigma^{-\eta n - p - 2(a_0+1)} \exp\left[-\frac{1}{2\sigma^2}(y^* - \mathbf{X}^*\beta)^T(y^* - \mathbf{X}^*\beta)\right] \exp\left[-\frac{b_0}{\sigma^2}\right] \\
&= \sigma^{-p} \exp\left[-\frac{1}{2\sigma^2}(\beta - \bar{\beta}_{n,\eta})^T \mathbf{X}^{*T}\mathbf{X}^*(\beta - \bar{\beta}_{n,\eta})\right] \\
&\quad \cdot \sigma^{-2(a_0+1)-\eta n} \exp\left[-\frac{1}{2\sigma^2}(y^{*T}y^* - \bar{\beta}_{n,\eta}^T \mathbf{X}^{*T}\mathbf{X}^*\bar{\beta}_{n,\eta}) - \frac{b_0}{\sigma^2}\right].
\end{aligned}$$

The first line of the final expression is a Gaussian density on β with mean $\bar{\beta}_{n,\eta}$ and variance matrix $\sigma^2\Sigma_{n,\eta}$, and the second line is Inv-gamma$(\sigma^2 \mid a_{n,\eta}, b_{n,\eta})$ with

$$a_{n,\eta} = a_0 + \eta n/2 \quad ; \quad b_{n,\eta} = b_0 + \frac{1}{2}\bar{\beta}_0^T \mathbf{X}_0^T \mathbf{X}_0 \bar{\beta}_0 + \frac{\eta}{2}(y^n)^T y^n - \frac{1}{2}\bar{\beta}_{n,\eta}^T \Sigma_{n,\eta}^{-1} \bar{\beta}_{n,\eta}.$$