The Berkson model: details

This note gives the missing details in the Berkson model.

The model is

$$y = \beta x + \epsilon$$

where y is the response, β the effect of the true covariate x with zero mean Gaussian noise ϵ. The issue is that x is not observed directly, but only through x_{obs}, where

$$x_{\text{obs}} = x + \nu$$

where ν is zero mean Gaussian noise. The parameters are: β has prior $\pi(\beta)$, x is apriori $\mathcal{N}(\mu_x I, \tau_x I)$, and τ_{obs} is the observation precision for x (ie $\text{Prec}(x_{\text{obs}}|x))$[1].

Assume that the precision of the observations y, τ_y, is known as it does not influence the calculations. Let $\theta = (\beta, \tau_x, \tau_{\text{obs}}, \mu_x)$. The full posterior is

$$\pi(x, \theta | y, x_{\text{obs}}) \;\propto\; \pi(\theta)\, \pi(x|\theta)\, \pi(x_{\text{obs}}|x, \theta)\, \pi(y|x, \theta)$$

Using that

$$\pi(x|\theta)\, \pi(x_{\text{obs}}|x, \theta) = \pi(x|x_{\text{obs}}, \theta)\, \pi(x_{\text{obs}}|\theta)$$

we get

$$\pi(x, \theta | y, x_{\text{obs}}) \;\propto\; \pi(\theta)\, \pi(x|x_{\text{obs}}, \theta)\, \pi(x_{\text{obs}}|\theta)\, \pi(y|x, \theta).$$

This means that x only enters in *one term* (apart from the likelihood) hence can be used as an ordinary latent model `f()`. Its easy to derive that

$$x|x_{\text{obs}}, \theta \;\sim\; \mathcal{N}\left(\frac{\tau_x \mu_x I + \tau_{\text{obs}} x_{\text{obs}}}{\tau_x + \tau_{\text{obs}}}, (\tau_x + \tau_{\text{obs}}) I\right).$$

and

$$x_{\text{obs}}|\theta \;\sim\; \mathcal{N}\left(\mu_x I, \frac{1}{1/\tau_x + 1/\tau_{\text{obs}}} I\right).$$

Note that $x_{\text{obs}}|\theta$ does not depend on x, hence conditionally on θ, its a constant. But it do need to be included in the model, as its log-density is

$$-\frac{n}{2}\log(2\pi) + \frac{n}{2}\log\left(\frac{1}{1/\tau_x + 1/\tau_{\text{obs}}}\right) - \frac{1}{2}\frac{1}{1/\tau_x + 1/\tau_{\text{obs}}}(x_{\text{obs}} - \mu_x I)^T (x_{\text{obs}} - \mu_x I)$$

and do depend on θ.

The last tweak, is that we do the change of variable from (x, β) to (u, β), where $u = \beta x$, so that

$$y = u + \epsilon.$$

This makes the implementation more convenient. Then we get

$$\begin{aligned}
\pi(u, \theta | y, x_{\text{obs}}) \;\;\propto\;\; & \pi(\theta) && (1)\\
& \pi(u|x_{\text{obs}}, \theta)\, \pi(x_{\text{obs}}|\theta) && (2)\\
& \pi(y|u, \theta). && (3)
\end{aligned}$$

where

$$u|\theta, x_{\text{obs}} \;\sim\; \mathcal{N}\left(\beta \frac{\tau_x \mu_x I + \tau_{\text{obs}} x_{\text{obs}}}{\tau_x + \tau_{\text{obs}}}, \frac{\tau_x + \tau_{\text{obs}}}{\beta^2} I\right).$$

[1] Note: The second argument in $\mathcal{N}(,)$ is the precision not the variance.