1 Priors on regression parameters

The posterior distribution for a standard 2D error-in-variables regression model, ignoring possible covariance of the errors, is (eg. eqn. 35 of D'Agostini et al. 2005):

$$
\begin{aligned}
p(\alpha, \beta, \sigma_y^2 | x, y) &= \prod_i (s_{y,i}^2)^{-1/2} \exp\left(-\frac{1}{2}\frac{(\beta x_i + \alpha - y_i)}{s_{y,i}^2}\right) p(\alpha, \beta, \sigma_y^2) && (1)\\
s_{y,i}^2 &= \sigma_y^2 + \sigma_{y,i}^2 + \beta^2 \sigma_{x,i}^2 && (2)
\end{aligned}
$$

Write the prior distribution

$$
p(\alpha, \beta, \sigma_y^2) = p(\alpha|\beta) p(\sigma_y^2|\beta) p(\beta) \tag{3}
$$

Let Θ be the angle between x-axis and the best fit line. Then $\beta = tan(\Theta)$ and the perpendicular offset of the best fit line from the origin is $\alpha_\perp = \alpha \cos\Theta$. As advocated by Hogg et al. 2010, we choose priors that are equivalent to uniform in Θ and $\alpha_\perp$. Given the change of variables rules for probability distributions (see eg. `https://onlinecourses.science.psu.edu/stat414/node/157`), we can determine the distribution for α and β.

If the probability density function of a random variable X is given as $f_X(x)$, it is possible to calculate the probability density function of some variable $Y = g(X)$. This is also called a "change of variable". If the function g is monotonic, then:

$$
f(y) = \left| \frac{d}{dy}(g^{-1}(y)) \right| \cdot f_X(g^{-1}(y)) \tag{4}
$$

where g^{-1} denotes the inverse function.

In our case, $\frac{d}{dx}(\tan^{-1}(\mathrm{x})) = \frac{1}{(1+x^2)}$. Therefore, a uniform distribution in Θ implies:

$$
p(\beta) = \frac{1}{\pi(1+\beta^2)} \tag{5}
$$

This is indeed the Student's t distribution, as stated by Andreon & Hurn 2012.

Now, regarding p($\alpha|\beta$),

$$
\begin{aligned}
\alpha &= \frac{\alpha_\perp}{\cos\Theta} && (6)\\
&= \frac{\alpha_\perp}{\cos(\arctan(\beta))} && (7)\\
&= \alpha_\perp (1+\beta^2)^{\frac{1}{2}} && (8)
\end{aligned}
$$

So therefore, a uniform distribution in $\alpha_\perp$ implies

$$
p(\alpha|\beta) \propto (1+\beta^2)^{\frac{1}{2}} \tag{9}
$$

Now, regarding p($\sigma_y^2|\beta$),

$$
\begin{aligned}
\sigma_y^2 &= \frac{\sigma_\perp^2}{\cos^2\Theta} && (10)\\
&= \frac{\sigma_\perp^2}{\cos^2(\arctan(\beta))} && (11)\\
&= \sigma_\perp^2 (1+\beta^2) && (12)
\end{aligned}
$$

So therefore, a uniform distribution in $\sigma_\perp^2$ implies

$$p(\sigma_y^2|\beta) \propto (1+\beta^2) \tag{13}$$

Putting this all together, we have

$$\begin{aligned} p(\alpha,\beta,\sigma_y^2) &= p(\alpha|\beta)p(\sigma_y^2|\beta)p(\beta) && (14)\\ &\propto (1+\beta^2)^{1/2} && (15) \end{aligned}$$

Then,

$$\begin{aligned} p(\alpha,\beta,\sigma_y^2|x,y) &= \frac{1}{2}\sum_i \left[\ln\left(\frac{1}{s_{y,i}^2}\right) - \frac{1}{2}\frac{(\beta x_i+\alpha-y_i)}{s_{y,i}^2} + \ln\left(1+\beta^2\right)\right] && (16)\\ & && (17) \end{aligned}$$

which agrees with equation 10 and 13 of Robotham and Obreschkow 2015:

$$p(\alpha,\beta,\sigma_y^2|x,y) = \frac{1}{2}\sum_i \left[\ln\left(\frac{(1+\alpha^2)}{s_{y,i}^2}\right) - \frac{1}{2}\frac{(\alpha x_i+\beta-y_i)}{s_{y,i}^2}\right] \tag{18}$$

It can be shown that, for a multivariate regression, the rotationally invariant priors are:

$$p(\alpha,\beta,\sigma_y^2) \propto (1+\sum_{j=1}^{D-1}\beta_j^2)^{1/2} \tag{19}$$

where D is the number of dimensions in the data.