Ridge-regression:

Priors: $\beta_i \sim N(b_i=0, S_i)$

$$\Rightarrow f(\beta_i \mid b_i=0, S_i) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\left(\frac{\beta_i - 0}{S_i}\right)^2}$$

$$= \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\left(\frac{\beta_i}{S_i}\right)^2}$$

Likelihoods: $y_i \sim N(x_i^T\beta, \sigma=1)$

$$\Rightarrow f(y_i \mid x_i^T\beta, \sigma=1) = \frac{1}{\sqrt{2\pi \cdot 1^2}} e^{-\frac{1}{2}\left(\frac{y_i - x_i^T\beta}{1^2}\right)^2}$$

$$= \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(y_i - x_i^T\beta)^2}$$

Posteriors: $P(\beta_i \mid y_i) \propto P(y_i \mid x_i^T\beta, \sigma=1)\, P(\beta_i \mid b_i=0, S_i)$

$$= \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(y_i - x_i^T\beta)^2} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\left(\frac{\beta_i}{S_i}\right)^2}$$

$$\propto \prod_{i=1}^{n} e^{-\frac{1}{2}(y_i - x_i^T\beta)^2} e^{-\frac{1}{2}\left(\frac{\beta_i}{S_i}\right)^2}$$

$$= e^{\sum_{i=1}^{n} -\frac{1}{2}(y_i - x_i^T\beta)^2 - \frac{1}{2}\left(\frac{\beta_i}{S_i}\right)^2}$$

"Bayesian do not optimize posterior distributions, they sample from them; but, the posterior distributions are nonetheless 'regularizations' of the likelihood through the prior.

log Posterior: $\log P(\beta_i \mid y_i) = \log e^{\sum_{i=1}^{n} -\frac{1}{2}(y_i - x_i^T\beta)^2 - \frac{1}{2}\left(\frac{\beta_i}{S_i}\right)^2}$

$$= \sum_{i=1}^{n} -\frac{1}{2}(y_i - x_i^T\beta)^2 - \frac{1}{2}\left(\frac{\beta_i}{S_i}\right)^2$$

$$\propto \sum_{i=1}^{n} \frac{1}{2}(y_i - x_i^T\beta)^2 + \frac{1}{2S_i^2} \sum_{i=1}^{n} \beta_i^2$$

$$= \sum_{i=1}^{n} \frac{1}{2}(y_i - x_i^T\beta)^2 + \lambda \sum_{j=1}^{n} \beta_j^2, \text{ let } \lambda = \frac{1}{2S_i^2}$$

Lasso Regression:

Priors: $\beta_i \sim \text{Laplace}(b_i=0, S_i)$

$$\Rightarrow f(\beta_i \mid b_i=0, S_i) = \frac{1}{2b} e^{-\frac{|\beta_i - 0|}{S_i}}$$

$$= \frac{1}{2b} e^{-\frac{|\beta_i|}{S_i}}$$

Likelihoods: $y_i \sim N(x_i^T\beta, \sigma=1)$

$$\Rightarrow f(y_i \mid x_i^T\beta, \sigma=1) = \frac{1}{\sqrt{2\pi \cdot 1^2}} e^{-\frac{1}{2}\left(\frac{y_i - x_i^T\beta}{1^2}\right)^2}$$

$$= \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(y_i - x_i^T\beta)^2}$$

Posteriors: $P(\beta_i \mid y_i) \propto P(y_i \mid x_i^T\beta, \sigma=1)\, P(\beta_i \mid b_i=0, S_i)$

$$= \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(y_i - x_i^T\beta)^2} \frac{1}{2b} e^{-\frac{|\beta_i|}{S_i}}$$

$$\propto \prod_{i=1}^{n} e^{-\frac{1}{2}(y_i - x_i^T\beta)^2} e^{-\frac{|\beta_i|}{S_i}}$$

$$= e^{\sum_{i=1}^{n} -\frac{1}{2}(y_i - x_i^T\beta)^2 - \frac{|\beta_i|}{S_i}}$$

log Posterior: $\log P(\beta_i \mid y_i) = \log e^{\sum_{i=1}^{n} -\frac{1}{2}(y_i - x_i^T\beta)^2 - \frac{|\beta_i|}{S_i}}$

$$= \sum_{i=1}^{n} -\frac{1}{2}(y_i - x_i^T\beta)^2 - \frac{1}{S_i}\sum_{i=1}^{n} |\beta_i|$$

$$\propto \sum_{i=1}^{n} \frac{1}{2}(y_i - x_i^T\beta)^2 + \frac{1}{S_i}\sum_{i=1}^{n} |\beta_i|$$

$$= \sum_{i=1}^{n} \frac{1}{2}(y_i - x_i^T\beta)^2 + \lambda \sum_{j=1}^{n} |\beta_j|, \text{ let } \lambda = \frac{1}{S_i}$$