Part 2:

when $a=1$

$$f(y|\beta) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi \cdot 1}} e^{-\frac{(y_i - x_i^T \beta_{p\times1})^2}{2\times1}}$$

$$\propto e^{-\frac{1}{2}\sum_{i=1}^{n}(y_i - x_i^T \beta_{p\times1})^2}$$

normal prior of $b_i = 0$

$$f(\beta|b_i) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}\,a} e^{-\frac{\beta_i^2}{2S_i^2}}$$

$$\propto e^{-\frac{1}{2}\sum_{i=1}^{n}\frac{\beta_i^2}{S_i^2}}$$

$$f(\beta|y) \propto f(y|\beta) f(\beta|b_i)$$

$$\propto \exp\left\{-\frac{1}{2}\left[\sum_{i=1}^{n}(y_i - x_i^T \beta_{p\times1})^2 + \sum_{i=1}^{n}\frac{\beta_i^2}{S_i^2}\right]\right\}$$

for $-\frac{1}{2}\left[\sum_{i=1}^{n}(y_i - x_i^T \beta_{p\times1})^2 + \sum_{i=1}^{n}\frac{\beta_i^2}{S_i^2}\right]$, the normal prior have analagous forms to the Lasso regression.

Laplace prior: $b_i = 0$.

$$f(\beta|b_i) = \frac{1}{2b} e^{-\frac{|\beta_i|}{S_i}}$$

$$\propto e^{-\frac{|\beta_i|}{S_i}}$$

$$f(\beta|y) \propto f(y|\beta) f(\beta|b_i)$$

$$\propto \exp\left\{-\frac{1}{2}\left[\sum_{i=1}^{n}(y_i - x_i^T \beta_{p\times1})^2\right] - \frac{|\beta_i|}{S_i}\right\}$$

for $-\frac{1}{2}\left[\sum_{i=1}^{n}(y_i - x_i^T \beta_{p\times1})^2\right] - \frac{|\beta_i|}{S_i})$. the Laplace prior have analagous forms the Ridge Regression

" Bayesians do not optimize posterior distributions. They sample from, but, the posterior distribution are nonetheless 'regularizations' of the likelihood through prior."

The posterior distribution is conclusion of prior and data.