Marco Reina 7066486
Samuele Serri 7069839

1. **a)**

   if $\lambda \to 0$ then the solution of (2.1) and (2.2) will be the same as minimizing RSS.
   if $\lambda \to \infty$ then the coefficients β_j for $j = 1, ..., p$ will be close to zero.

   b)

   Lasso regularization has the property of setting some coefficients exactly to zero with a λ large enough, therefore if our goal is to do variable selection we should use L_1 regularization.

2. **a)**

$$f(Y|X, \beta) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(- \frac{|x_i|}{\sigma_2} \right)$$

   b)

$$p(\beta) = \prod_{j=1}^{p} \frac{1}{2b} \exp\left(- \frac{|\beta_j|}{b} \right)$$

   The posterior distribution takes the form:

$$p(\beta|X, Y) \propto f(Y|X, \beta)p(\beta|X) = f(Y|X, \beta)p(\beta) \quad [1]$$

$$p(\beta|X,Y) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(- \frac{|x_i|}{\sigma_2} \right) \prod_{j=1}^{p} \frac{1}{2b} \exp\left(- \frac{|\beta_j|}{b} \right) =$$

$$= \prod_{i=1}^{n} \prod_{j=1}^{p} \frac{1}{\sqrt{2\pi\sigma^2}2b} \exp\left(\frac{-|x_i|}{\sigma^2} + \frac{-|\beta_j|}{b} \right)$$

c)

[1] The formula for the posterior distribution was taken from the book *James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). An introduction to statistical learning* p. 251