Q20

Rahul Atre

2023-10-20

Q20 [Regularization]

We will now derive the Bayesian connection to the LASSO and ridge regression discussed class.

1. Suppose that $y_i = \beta 0 + \sum_{i=1}^{p} x_{ij}\beta_j + \epsilon_i$ where $\epsilon_1, \epsilon_2, ..., \epsilon_n$ are independent and identically distributed from a $N(0, \sigma^2)$ distribution. Write out the likelihood for the data.

2. Assume the following prior for $\beta : \beta_1, ..., \beta_p$ are independent and identically distributed according to a double-exponential distribution with mean 0 and common scale parameter b: i.e. $p(\beta) = \frac{1}{2b}exp(-|\beta|/b)$. Write out the posterior for β in this 2b setting.

3. Argue that the LASSO estimate is the mode for β under this posterior distribution.

4. Now assume the following prior for $\beta : \beta_1, ..., \beta_p$ are independent and identically distributed according to a normal distribution with mean zero and variance c. Write out the posterior for β in this setting.

5. Argue that the ridge regression estimate is both the mode and the mean for β under this posterior distribution.