EML Assignment 3 Problem 3

Muhammed Saeed and Ali Salaheldin Ali Ahmed

December 2022

Question 1

1 The ridge regression optimization problem in this setting is given by:

$$\min_{\beta_1,\beta_2} \left((y_1 - \beta_1 x_{11} - \beta_2 x_{12})^2 + (y_2 - \beta_1 x_{21} - \beta_2 x_{22})^2 + \lambda(\beta_1^2 + \beta_2^2)\right)$$

Where λ is the regularization parameter.

2 The solution space for ordinary least squares regression in this setting is given by a line with slope -1.

3 We can show that the ridge coefficient estimates satisfy $\beta_1 = \beta_2$ as follows:

From the given constraints, we have $y_1 = -y_2$, $x_{11} = -x_{21}$, $x_{12} = -x_{22}$ and $x_{11} = x_{12}$. Substituting these into the ridge regression optimization problem, we get:

$$\begin{aligned}
&\min_{\beta_1,\beta_2} \left((y_1 - x_{11}(\beta_1 + \beta_2))^2 + (x_{11}(\beta_1 + \beta_2) - y_1)^2 + \lambda(\beta_1^2 + \beta_2^2)\right) \\
=&\min_{\beta_1,\beta_2} \left(2(x_{11}(\beta_1 + \beta_2) - y_1)^2 + \lambda(\beta_1^2 + \beta_2^2)\right) \\
=&\min_{\beta_1,\beta_2} \left(2y_1^2 - 4x_{11}y_1(\beta_1 + \beta_2) + 2x_{11}^2(\beta_1 + \beta_2)^2 + \lambda(\beta_1^2 + \beta_2^2)\right)
\end{aligned}$$

To minimize this expression, we can set the partial derivatives of each of the terms with respect to β_1 and β_2 to zero, which gives us the following equations:

$$\begin{gathered}
\frac{\partial}{\partial \beta_1} \left(2y_1^2 - 4x_{11}y_1(\beta_1 + \beta_2) + 2x_{11}^2(\beta_1 + \beta_2)^2 + \lambda(\beta_1^2 + \beta_2^2)\right) = 0 \\
-4x_{11}y_1 + 4x_{11}^2(\beta 1 + \beta_2) + 2\lambda\beta_1 = 0 \\
\frac{\partial}{\partial \beta_2} \left(2y_1^2 - 4x_{11}y_1(\beta_1 + \beta_2) + 2x_{11}^2(\beta_1 + \beta_2)^2 + \lambda(\beta_1^2 + \beta_2^2)\right) = 0 \\
-4x_{11}y_1 + 4x_{11}^2(\beta 1 + \beta_2) + 2\lambda\beta_2 = 0
\end{gathered}$$

And by subtracting the two simplified equations from above, we get:

$$2\lambda(\beta_1 - \beta_2) = 0$$
$$\beta_1 = \beta_2$$

4 In the previous example, we have shown that in the special case where $n = 2$, $p = 2$, $x_{11} = x_{12}$, $x_{21} = x_{22}$, $y_1 + y_2 = 0$, $x_{11} + x_{21} = 0$, and $x_{12} + x_{22} = 0$, the coefficient estimates for ridge regression satisfy $\beta_1 = \beta_2$. These conditions mean that the predictor variables X_1 and X_2 have exactly the same values across the two observations.

This shows that when the variables are correlated and the response variable is a linear combination of the predictor variables, ridge regression tends to give similar coefficient values to the correlated variables.