Authors: Wojciech Ptaś (7042843), Mhd Jawad Al Rahwanji (703890)

Problem 3

1.

Considering our assumptions:

- $\beta_0 = 0$
- $y_1 + y_2 = 0 \iff y_2 = -y_1$
- $x_{11} = x_{12}$
- $x_{21} = x_{22}$
- $x_{11} + x_{21} = 0 \iff x_{21} = -x_{11}$

In this setting we need to minimize the following value:

$RSS + \lambda \sum_{j=1}^{p} \beta_j^2 =$

$= \sum_{i=1}^{n} (y_i - \beta_0 - \sum_{j=1}^{p} \beta_j x_{ij})^2 + \lambda \sum_{j=1}^{p} \beta_j^2 =$

Because $n = 2$ and $p = 2$, we can expand the sums and substitute some of the values

$= (y_1 - \beta_1 x_{11} - \beta_2 x_{11})^2 + (-y_1 + \beta_1 x_{11} + \beta_2 x_{11})^2 + \lambda\beta_1^2 + \lambda\beta_2^2 =$

$= 2(y_1 - \beta_1 x_{11} - \beta_2 x_{11})^2 + \lambda\beta_1^2 + \lambda\beta_2^2$

In order to calculate the value of coefficients, we need to minimize the expression above

2.

In order to find the solution space for least squares regression we have to find the values that minimizes the RSS. In this particular case, it can be evaluated as:

$RSS = 2(y_1 - \beta_1 x_{11} - \beta_2 x_{11})^2$

Now we need to calculate the derivatives:

- $\frac{\partial RSS}{\partial \beta_1} = -4x_{11}(y_1 - \beta_1 x_{11} - \beta_2 x_{11})$
- $\frac{\partial RSS}{\partial \beta_2} = -4x_{11}(y_1 - \beta_1 x_{11} - \beta_2 x_{11})$

We can observe that both derivaties are in fact equal. To find the value of coefficients we need to solve the equations above for 0, which yields a linear equation. The solution space for linear equation with two variables is a straight line, so this is how our solution space looks like

3.

To show that in this example, ridge coefficient estimates satisfy $\hat{\beta}_1 = \hat{\beta}_2$ we need to calculate the derivative of following expression: For the sake of simplicity, we won't use the "hat" symbols in the proof.

$f(\hat{\beta}_1, \hat{\beta}_2) = 2(y_1 - \beta_1 x_{11} - \beta_2 x_{11})^2 + \lambda\beta_1^2 + \lambda\beta_2^2$

$\frac{\partial f}{\partial \beta_1} = -4x_{11}(y_1 - \beta_1 x_{11} - \beta_2 x_{11}) + 2\lambda\beta_1$

$\frac{\partial f}{\partial \beta_2} = -4x_{11}(y_1 - \beta_1 x_{11} - \beta_2 x_{11}) + 2\lambda\beta_2$

Two find the value of coefficients, we now need to solve following system of equations:

$\frac{\partial f}{\partial \beta_1} = 0$

$\frac{\partial f}{\partial \beta_2} = 0$

First, let us simplify the equations

$-4x_{11}(y_1 - \beta_1 x_{11} - \beta_2 x_{11}) + 2\lambda\beta_1 = 0$

$-4x_{11}y_1 + 4x_{11}^2\beta_1 + 4x_{11}^2\beta_2 + 2\lambda\beta_1 = 0$

$\beta_1(4x_{11}^2 + 2\lambda) + 4x_{11}^2\beta_2 = 4x_{11}y_1$

Analogically we proceed with the second equation, as a result we get:

$\beta_1(4x_{11}^2 + 2\lambda) + 4x_{11}^2\beta_2 = 4x_{11}y_1$

$4x_{11}^2\beta_1 + (4x_{11}^2 + 2\lambda)\beta_2 = 4x_{11}y_1$

Now, solving for β_1 and β_2

To simplify thigs: - $a = 4x_{11}^2$
- $b = 4x_{11}^2 + 2\lambda$
- $c = 4x_{11}y_1$

$b\beta_1 + a\beta_2 = c$

$a\beta_1 + b\beta_2 = c$

$\beta_1 = \frac{c-a\beta_2}{b}$

$a\frac{c-a\beta_2}{b} + b\beta_2 = c$

$\frac{ac}{b} - \frac{a^2}{b}\beta_2 + b\beta_2 = c$

$\beta_2(\frac{b^2-a}{b}) = \frac{bc-ac}{b}$

$\beta_2 = \frac{bc-ac}{b}\frac{b}{b^2-a} = \frac{c(b-a)}{(b-a)(b+a)} = \frac{c}{a+b}$

$\beta_1 = \frac{c-a\frac{c}{a+b}}{b} = \frac{\frac{ac+bc-ac}{a+b}}{b} = \frac{\frac{bc}{a+b}}{b} = \frac{c}{a+b}$

This means that $\beta_1 = \beta_2$, quod erat demonstrantum

4.

In this setting, the both features had the same value for every sample and the estimated coefficients were also equal. It corresponds to the fact, that in ridge regression, correlated variables (i. e equal especially) will have very similar coefficients.