Date __/__/__
Page ____
myCOMPANION

Divyatmika Inu

HW4INF552

6.8.5

(a.) According to the setting ($x_{11} = x_{12} = x_1$) and ($x_{21} = x_{22} = x_2$), the ridge ~~comp~~ regression problem seeks to minimize

$$(y_1 - \hat{\beta}_1 x_1 - \hat{\beta}_2 x_2)^2 + (y_2 - \hat{\beta}_1 x_2 - \hat{\beta}_2 x_2)^2 + \lambda(\hat{\beta}_1^2 + \hat{\beta}_2^2)$$

(b.) By taking the derivative of the above expression with respect to $\hat{\beta}_1$ and $\hat{\beta}_2$, & setting them equal to 0, we obtain

$$\hat{\beta}_1 (x_1^2 + x_2^2 + \lambda) + \hat{\beta}_2 (x_1^2 + x_2^2) = y_1 x_1 + y_2 x_2$$

and

$$\hat{\beta}_1 (x_1^2 + x_2^2) + \hat{\beta}_2 (x_1^2 + x_2^2 + \lambda) = y_1 x_1 + y_2 x_2$$

By substracting the two expressions above we get

$$\underline{\hat{\beta}_1 = \hat{\beta}_2}$$

(c.) According to the setting ($x_{11} = x_{12} = x_1$ & $x_{21} = x_{22} = x_2$), the lasso optimization problem seeks to minimize :→

$$(y_1 - \hat{\beta}_1 x_1 - \beta_2 x_1)^2 + (y_2 - \hat{\beta}_1 x_2 - \hat{\beta}_2 x_2)^2 + \lambda(|\hat{\beta}_1| + |\hat{\beta}_2|)$$

(d.) Using the alternate form of lasso problem :-

$$(y_1 - \hat{\beta}_1 x_1 - \hat{\beta}_2 x_1)^2 + (y_2 - \hat{\beta}_1 x_2 - \hat{\beta}_2 x_2)^2 \text{ subject to } |\hat{\beta}_1| + |\hat{\beta}_2| \leq s$$

Geometrically the lasso constraint take the form of diamond centered at origin of the plane ($\hat{\beta}_1, \hat{\beta}_2$) which intersects the axes at a distances from the origin.

By using the setting of this problem ($x_{11} = x_{12} = x_1, x_{21} = x_{22} = x_2$, $x_1 + x_2 = 0$ and $y_1 + y_2 = 0$) we have to minimize

$$2[y_1 - (\hat{\beta}_1 + \hat{\beta}_2)x_1]^2 \geq 0$$

This optimization has a simple solution: $\hat{\beta_1} + \hat{\beta_2} = y_1/x_1$. Geometrically, this is a line parallel to the edge of the diamond of the constraints.

Now, solutions to the lasso optimization problems are contours of f^n $[y_1 - (\hat{\beta_1} + \hat{\beta_2})x_1]^2$ that intersects the diamond of the constraints.

So the entire edge $\hat{\beta_1} + \hat{\beta_2} = s$ (as is the edge $\hat{\beta_1} + \hat{\beta_2} = -s$) is a potential solution to the lasso optimization. Thus the

$$\left\{ (\hat{\beta_1}, \hat{\beta_2}) : \hat{\beta_1} + \hat{\beta_2} = s \text{ with } \hat{\beta_1}, \hat{\beta_2} \geq 0 \text{ and } \hat{\beta_1} + \hat{\beta_2} = -s \text{ ; with } \hat{\beta_1}, \hat{\beta_2} \leq 0 \right\}$$