HW4_Myunghee_ID_2446752777

4. ISLR, 6.8.5

	p_1	p_2	y
n_1	x_{11}	$x_{12} = x_{11}$	y_1
n_2	$x_{21} = -x_{11}$	$x_{22} = -x_{11}$	$y_2 = -y_1$

Thus, it can be simplified as below.

	p_1	p_2	Y
n_1	x	x	y
n_2	-x	-x	-y

(a) Write out the ridge regression optimization problem in this setting.

Minimize:

$$f(\hat{\beta}_1, \hat{\beta}_2) = (y - \hat{\beta}_1 x - \hat{\beta}_2 x)^2 + (-y + \hat{\beta}_1 x + \hat{\beta}_2 x)^2 + \lambda\left(\hat{\beta}_1{}^2 + \hat{\beta}_2{}^2\right)$$
$$= 2(y - \hat{\beta}_1 x - \hat{\beta}_2 x)^2 + \lambda\left(\hat{\beta}_1{}^2 + \hat{\beta}_2{}^2\right)$$

(b) Argue that in this setting, the ridge coefficient estimates satisfy ˆβ1 = ˆβ2.

$$\frac{\partial f}{\partial \hat{\beta}_1} = 4(y - \hat{\beta}_1 x - \hat{\beta}_2 x)(-x) + 2\lambda\hat{\beta}_1 = -4xy + 4\hat{\beta}_1 x^2 + 4\hat{\beta}_2 x^2 + 2\lambda\hat{\beta}_1 = 0$$

$$(2x^2 + \lambda)\hat{\beta}_1 = 2xy - 2\hat{\beta}_2 x^2$$

$$\hat{\beta}_1 = \frac{2xy - 2\hat{\beta}_2 x^2}{2x^2 + \lambda}$$

$$\frac{\partial f}{\partial \hat{\beta}_2} = 4(y - \hat{\beta}_1 x - \hat{\beta}_2 x)(-x) + 2\lambda\hat{\beta}_2 = -4xy + 4\hat{\beta}_1 x^2 + 4\hat{\beta}_2 x^2 + 2\lambda\hat{\beta}_2 = 0$$

$$(2x^2 + \lambda)\hat{\beta}_2 = 2xy - 2\hat{\beta}_1 x^2$$

$$\hat{\beta}_2 = \frac{2xy - 2\hat{\beta}_1 x^2}{2x^2 + \lambda}$$

$\hat{\beta}_1$ and $\hat{\beta}_2$ are symmetrical and seem to satisfy $\hat{\beta}_1 = \hat{\beta}_2$.

(c) Write out the lasso optimization problem in this setting.

Minimize:

$$f(\hat{\beta}_1, \hat{\beta}_2) = (y - \hat{\beta}_1 x - \hat{\beta}_2 x)^2 + (-y + \hat{\beta}_1 x + \hat{\beta}_2 x)^2 + \lambda(|\hat{\beta}_1| + |\hat{\beta}_2|)$$
$$= 2(y - \hat{\beta}_1 x - \hat{\beta}_2 x)^2 + \lambda(|\hat{\beta}_1| + |\hat{\beta}_2|)$$

(d) Argue that in this setting, the lasso coefficients ˆβ1 and ˆβ2 are not unique—in other words, there are many possible solutions to the optimization problem in (c). Describe these solutions.

If $\hat{\beta}_1 \geq 0$,

$$\frac{\partial f}{\partial \hat{\beta}_1} = 4(y - \hat{\beta}_1 x - \hat{\beta}_2 x)(-x) + \lambda = -4xy + 4\hat{\beta}_1 x^2 + 4\hat{\beta}_2 x^2 + \lambda = 0$$

$$2x^2 \hat{\beta}_1 = 2xy - 2\hat{\beta}_2 x^2 - \lambda$$

$$\hat{\beta}_1 = \frac{2xy - 2\hat{\beta}_2 x^2 - \lambda}{2x^2}$$

If $\hat{\beta}_1 < 0$,

$$\frac{\partial f}{\partial \hat{\beta}_1} = -4xy + 4\hat{\beta}_1 x^2 + 4\hat{\beta}_2 x^2 - \lambda = 0$$

$$2x^2 \hat{\beta}_1 = 2xy - 2\hat{\beta}_2 x^2 + \lambda$$

$$\hat{\beta}_1 = \frac{2xy - 2\hat{\beta}_2 x^2 + \lambda}{2x^2}$$

If $\hat{\beta}_2 \geq 0$,

$$\frac{\partial f}{\partial \hat{\beta}_2} = 4(y - \hat{\beta}_1 x - \hat{\beta}_2 x)(-x) + \lambda = -4xy + 4\hat{\beta}_1 x^2 + 4\hat{\beta}_2 x^2 + \lambda = 0$$

$$2x^2 \hat{\beta}_2 = 2xy - 2\hat{\beta}_1 x^2 - \lambda$$

$$\hat{\beta}_2 = \frac{2xy - 2\hat{\beta}_1 x^2 - \lambda}{2x^2}$$

If $\hat{\beta}_2 < 0$,

$$\frac{\partial f}{\partial \hat{\beta}_2} = -4xy + 4\hat{\beta}_1 x^2 + 4\hat{\beta}_2 x^2 - \lambda = 0$$

$$2x^2 \hat{\beta}_2 = 2xy - 2\hat{\beta}_1 x^2 + \lambda$$

$$\hat{\beta}_2 = \frac{2xy - 2\hat{\beta}_1 x^2 + \lambda}{2x^2}$$

According to signs of $\hat{\beta}_1$ and $\hat{\beta}_2$, there are four combinations of $\hat{\beta}_1$ and $\hat{\beta}_2$. For example, if $\hat{\beta}_1 \geq 0$ and $\hat{\beta}_2 < 0$, $\hat{\beta}_1$ and $\hat{\beta}_2$ are quite different.

$$\hat{\beta}_1 = \frac{2xy - 2\hat{\beta}_2 x^2 - \lambda}{2x^2}$$

$$\hat{\beta}_2 = \frac{2xy - 2\hat{\beta}_1 x^2 + \lambda}{2x^2}$$