Supplement for Ridge regression

Tae Geun Kim

2021-05-14

Contents

Additional Knowledge for Ridge

Intercept for centered input

Let's start with an exercise.

Exercise 3.5 in ESL

Show that the ridge regression problem is equivalent to the problem :

$$\hat{\beta}^c = \underset{\beta^c}{\text{argmin}} \left[\sum_{i=1}^N \left\{ y_i - \beta_0^c - \sum_{j=1}^p (x_{ij} - \bar{x}_j)\beta_j^c \right\}^2 + \lambda \sum_{j=1}^p (\beta_j^c)^2 \right]$$

Answer for Exercise 3.5

Let consider next transformations. Then that's the answer.

$$\beta_0 \rightarrow \beta_0^c - \sum_{j=1}^p \bar{x}_j \beta_j, \quad \beta_j \rightarrow \beta_j^c \ (j > 0)$$

From this exercise, we can see the alternative form of the ridge regression. If we take the form, than it's convenient to estimate $\hat{\beta}_0$.

Estimate $\hat{\beta}_0$

Let denote the loss function.

$$L(\beta_0^c) \equiv \left[\sum_{i=1}^N \left\{ y_i - \beta_0^c - \sum_{j=1}^p (x_{ij} - \bar{x}_j)\beta_j^c \right\}^2 + \lambda \sum_{j=1}^p (\beta_j^c)^2 \right]$$

Then to calculate $\hat{\beta}_0^c$, we need differentiations.

$$\begin{aligned}
\frac{\partial}{\partial \beta_0^c} L(\beta_0^c) &= -2 \sum_{i=1}^N \left[y_i - \beta_0^c - \sum_{j=1}^p (x_{ij} - \bar{x}_j)\beta_j^c \right] = 0 \\
\Rightarrow \ N\hat{\beta}_0^c &= \sum_{i=1}^N \left(y_i - \sum_{j=1}^p (x_{ij} - \bar{x}_j)\beta_j^c \right) \qquad (1) \\
\therefore \hat{\beta}_0^c &= \frac{1}{N} \sum_{i=1}^N y_i = \bar{y}
\end{aligned}$$

Effective Degree of Freedom

Degree of freedom of OLS

OLS is written as next form.

$$\hat{\mathbf{y}} = \mathbf{H}\mathbf{y}, \quad \text{where } \mathbf{H} = \mathbf{X}(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T$$