

CS385 Machine Learning Homework 8

Zhou Litao 518030910407 F1803016

April 28, 2021, Spring Semester

Exercise 1 Please prove that the mean value of beta given X and Y ($\Pr[\beta|Y, X]$) on Page 14 equals to the solution to the ridge regression. This is taught in the class, not in the note.

Solution.

The posterior distribution for β for the linear version of bayesian regression is

$$\Pr[\beta | Y, X] = N \left(\tau^2 X^T (\tau^2 X X^T + \sigma^2 I_n)^{-1} Y, \tau^2 I_p - \tau^2 X^T (\tau^2 X X^T + \sigma^2 I_n)^{-1} \tau^2 X \right)$$

We can consider the mean value to be the optimal solution that the ridge regression model should give. We first set the λ of the ridge regression to be $\frac{\sigma^2}{\tau^2}$, we show that the mean value is equal to the solution of ridge regression.

$$\begin{aligned} \tau^2 X^T (\tau^2 X X^T + \sigma^2 I_n)^{-1} Y &= X^T (X X^T + \lambda I_n)^{-1} Y \\ &= (X^T X + \lambda I_n)^{-1} (X^T X + \lambda I_n) X^T (X X^T + \lambda I_n)^{-1} Y \\ &= (X^T X + \lambda I_n)^{-1} X^T (X X^T + \lambda I_n) (X X^T + \lambda I_n)^{-1} Y \\ &= (X^T X + \lambda I_n)^{-1} X^T Y \end{aligned} \tag{1}$$

□