

Machine Learning

ASSIGNMENT 6

1. Use multilinear regression analysis on the following set of three points. We use the notation in the handout on multilinear regression analysis. In this problem, $p = 2$, and $n = 3$.

$$\begin{aligned}x_1 &= (1, 2), \quad y_1 = 1 \\x_2 &= (-3, 2), \quad y_2 = 2 \\x_3 &= (2, -1), \quad y_3 = -1\end{aligned}$$

Estimate the parameter vector $\beta = [\beta_0 \quad \beta_1 \quad \beta_2]^T$. Do not use any computer package or program to do this problem.

Hint:

$$\hat{\beta} = [1/12 \quad -1/4 \quad 7/12]^T$$

2. There are two parts in this problem. In this problem, $p = 2$, and $n = 3$.

- (a) Try doing multilinear regression analysis on

$$\begin{aligned}x_1 &= (1, 2), \quad y_1 = 1 \\x_2 &= (2, 4), \quad y_2 = 2 \\x_3 &= (-3, -6), \quad y_3 = -3\end{aligned}$$

Note what happens.

- (b) For the data in part (a) of the problem, perform ridge regression. Tabulate results for:

$$\lambda = \pm 10^{-5}, \pm 10^{-4}, \pm 10^{-3}, \pm 10^{-2}, \pm 10^{-1}$$

In each case record: the β -vector, the corresponding value of minimized error E , and of course λ . Note that

$$E = \|Y - X\beta\|^2$$

Also plot on a log-log scale E versus $|\lambda|$.

There should be two curves: one for positive λ 's, and one for negative λ 's.