

Office hours:

Office Hours March 7-10:

- Tuesday 10-11:30am, rm. 1021 in Communications Building, Lichuan Deng
- Tuesday 2:00-3:50, Atanasoff 213 or 113, Prof. Quinn
- Wednesday 12-1:30, Atanasoff 213 or 113, Prof. Quinn
- Wednesday 2:10 - 3:40 pm rm. 1021 in Communications Building, Scott Song
- Thursday 10-11:30am, rm. 1021 in Communications Building, Lichuan Deng

1. Could we go over homework 3 or the concepts in that homework? Specifically about under/overfitting and how test/train mse's relate to the true $f(x)$ mse's.
2. What is the relationship between AIC (or BIC) with Lasso (or ridge regression)?

Could you please go over the main difference between lasso and ridge regression?

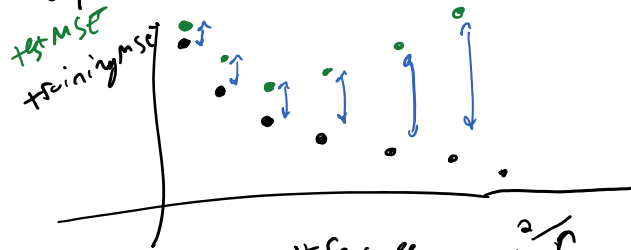
- HW 4
- Exam from previous semester

$$\hat{y} = \beta_0 + \beta_1 x_1 + \dots + \beta_p x_p$$

$$\hat{y} = \beta_0 + \beta_1 x_1 + \dots + \beta_p x_p$$

$\hat{y} = \beta_0^*$ typically underfit

$\hat{y} = \beta_0^{OLS} + \beta_1^{OLS} x_1 + \dots + \beta_p^{OLS} x_p$ typically overfit



we have formulas $\lambda_{AIC} = \frac{2}{n}$, $\lambda_{BIC} = \frac{\log n}{n}$

$$\min \text{training MSE} + \lambda \sum_{j=1}^p |\beta_j|^p$$

features (non-convex hard to solve)

Subset searches (exhaustive, greedy)

$$\min \text{training MSE} + \lambda \sum_{j=1}^p |\beta_j|^1$$

$$\min \text{training MSE} + \lambda \sum_{j=1}^p |\beta_j|^2$$

Lasso
(some feature selection)
ridge

both convex
don't do feature selection like AIC/BIC

