CIS 4190/5190: Applied Machine Learning — Spring 2023

**Homework 2**

*Handed Out: January 25* — *Due: February 8*

**Name:** Qihang Dai

**PennKey:** ahgdyyycc

**PennID:** 78803164

Note: This document is a read-only file. To create an editable version click on Menu in the top left corner of your screen and choose the Copy Project option.

# 1 Multiple Choice & Written Questions

1. (a) decrease bias and increase variance
   (b) increase bias and decrease variance
   (c) decrease bias and increase variance

   This means the modle is overfitting: low bias and high variance. thus we want a more generalized modle and a more simple prediciton. Thus we increase n, decrease $\lambda$, and decrease d.

2. (A) global maxium

3. (a)
$$\frac{\partial R_1}{\partial B_j} = \lambda sgn(B_j)$$
$$\frac{\partial R_2}{\partial B_j} = 2\lambda B_j^2$$
   (b) since L1 regularization's gradient is independent of $w$, it will do a better job to push weight to zero when weight is small.

4. (a) since x > 0, we have a = 1 and b = 0. y = x

for x distribution on [-1, 0], we have:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(x_i - 0)^2 \tag{1}$$

$$= \frac{1}{n}\sum_{i=1}^{n} x_i^2 \tag{2}$$

$$= \int_{-1}^{0} x^2\, dx \tag{3}$$

$$= \frac{1}{3} \tag{4}$$

(b) the learned model should be y = 0 (a = 0 and b = 0). for MSE on X U[0, 1], we have: y = x and MSE is still $\frac{1}{3}$

5.

$$f_{\hat{\beta}}(x) = \hat{\beta}^T x = x^T \hat{\beta}$$
$$\hat{\beta} = (X^T X)^{-1} X^T Y$$
$$f_{\hat{\beta}}(x) = x^T (X^T X)^{-1} X^T Y$$
$$Y = (y1, y2, ..., yn)^T, wehave$$
$$f_{\hat{\beta}}(x) = x^T (X^T X)^{-1} X^T (y1, y2, ..., yn)^T$$
$$= \sum_{i=1}^{n} x^T (X^T X)^{-1} X^T yi$$
$$k_i = x^T (X^T X)^{-1} X^T I_i$$

$I_i$ represent (d x 1) vector where only ith element is 1 and others are 0.

# 2 Python Programming Questions

TODO: Please add the resulting plots for Question 1.3 and Question 1.4.

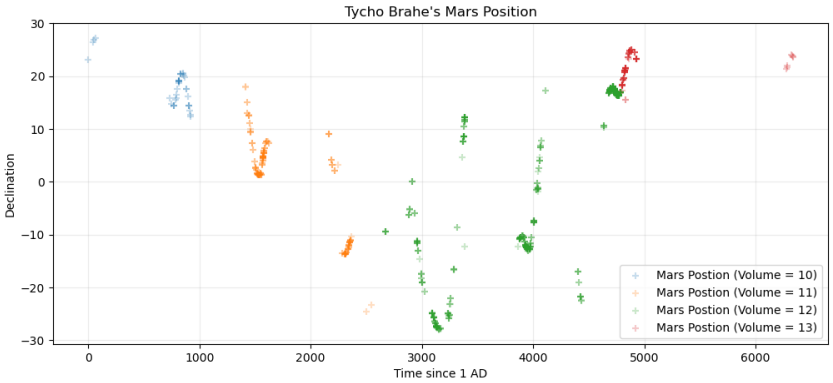


Figure 1: Figure