STATS 202: Statistical Learning and Data Science

Instructor: Linh Tran

HOMEWORK # 1

Due date: July 3, 2025

Stanford University

Introduction

Homework problems are selected from the course textbook: *An Introduction to Statistical Learning*.

Problem 1 (4 points)

Chapter 2, Exercise 3 (p. 53).

Problem 2 (4 points)

Chapter 2, Exercise 7 (p. 54).

Problem 3 (5 points)

Consider a linear regression model with p parameters, which is fit with a least squares estimator to a set of training data $(x_1, y_1), \ldots, (x_n, y_n) : (x_i, y_i) \overset{iid}{\sim} P_0$. Let $n > p$ and $\hat{\beta}$ be the least squares estimate. Suppose we have some test data $(\tilde{x}_1, \tilde{y}_1), \ldots, (\tilde{x}_m, \tilde{y}_m) : (\tilde{x}_i, \tilde{y}_i) \overset{iid}{\sim} P_0$. If $\mathcal{R}_{tr}(\beta) = \frac{1}{n} \sum_{i=1}^{n} (y_i - \beta^T x_i)^2$ and $\mathcal{R}_{te}(\beta) = \frac{1}{m} \sum_{i=1}^{m} (\tilde{y}_i - \beta^T \tilde{x}_i)^2$, prove that

$$\mathbb{E}\left(\mathcal{R}_{tr}(\hat{\beta})\right) \leq \mathbb{E}\left(\mathcal{R}_{te}(\hat{\beta})\right) \tag{1}$$

Problem 4 (5 points)

Chapter 12, Exercise 1 (p. 548).

Problem 5 (5 points)

Chapter 12, Exercise 2 (p. 548).

Problem 6 (4 points)

Chapter 12, Exercise 4 (p. 549).

Problem 7 (4 points)

Chapter 12, Exercise 9 (p. 550).

Problem 8 (4 points)

Chapter 3, Exercise 4 (p. 122).

Problem 9 (5 points)

Chapter 3, Exercise 9 (p. 123). In parts (e) and (f), you need only try a few interactions and transformations.

Problem 10 (5 points)

Chapter 3, Exercise 14 (p. 127).

Problem 11 (5 points)

Let $x_1, \ldots, x_n$ be a fixed set of input points and $y_i = f(x_i) + \epsilon_i$, where $\epsilon_i \overset{iid}{\sim} P_\epsilon$ with $\mathbb{E}(\epsilon_i) = 0$ and $\text{Var}(\epsilon_i) < \infty$. Prove that the MSE of a regression estimate $\hat{f}$ fit to $(x_1, y_1), \ldots, (x_n, y_n)$ for a <u>random</u> test point x_0 or $\mathbb{E}\left(y_0 - \hat{f}(x_0)\right)^2$ decomposes into variance, square bias, and irreducible error components. *Hint: You can apply the bias-variance decomposition proved in class.*

Problem 12 (Bonus 5 points)

Consider the regression through the origin model (i.e. with no intercept):

$$y_i = \beta x_i + \epsilon_i \tag{2}$$

(a) *(1 point)* Find the least squares estimate for β.

(b) *(2 points)* Assume $\epsilon_i \overset{iid}{\sim} P_\epsilon$ such that $\mathbb{E}(\epsilon_i) = 0$ and $\text{Var}(\epsilon_i) = \sigma^2 < \infty$. Find the standard error of the estimate.

(c) *(2 points)* Find conditions that guarantee that the estimator is consistent. *n.b. An estimator $\hat{\beta}_n$ of a parameter β is consistent if $\hat{\beta} \xrightarrow{p} \beta$, i.e. if the estimator converges to the parameter value in probability.*