

## CMSE 381: HW5

- 1 (Challenging Question not required) Exercise 4.7.4 (Textbook ISLR Chapter 4.7, Question 4)
- 2 (20 pts) Exercise 4.7.5
- 3 (10 pts) Exercise 4.7.6
- 4 (10 pts) Exercise 4.7.8
- 5 (20 pts) For a classification problem with  $K = 2$  ( $Y \in \{0, 1\}$ ), we know the oracle classifier is

$$C(x) = j, \text{ if } p_j(x) = \max \{p_0(x), p_1(x)\},$$

which is based on the loss with equal weight for Type I and II error. If we know Type I error will cost \$1000 while Type II error will cost \$3000. Derive the new oracle classifier which minimizes this cost.

- 6 (20 pts) Exercise 4.7.13
- 7 (20 pts) Use the wine.csv data and HW2.R data as in HW2 to perform LDA.
  - a. Generate a confusion table using the test data.
  - b. It will cost \$3000 if we classify class1 wine as class2; while it will cost \$ 1000 if we classify class2 wine as class1. Build a modified LDA according to Q5.
  - c. Calculate the cost on test data using ordinary LDA and the modified LDA respectively.
- 8 (Challenging Question not require) For linear regression, prove that the LOOCV solution has a close form as in the book.