MATH569 Homework 3 (total=40 pts)

Due March 6th Instructor: Lulu Kang

1. (10 pts) Ex. 3.12 from the textbook.

2. (10 pts) Ex 3.30. from the textbook.

   Hints of Problem 1 & 2: Consider the ridge regression $\mathbf{y} = \mathbf{X}\beta + \epsilon$, which is to minimize the least square plus the l_2 penalty. You can rewrite the objective function (Equation 3.41 from the textbook) into the following slightly changed version.

   $$\min_{\beta} ||\mathbf{y} - \mathbf{X}\beta||^2 + \gamma^2 ||\beta||_2^2$$

   Now if we augment the original regression input and output into following format:

   $$\tilde{X} = \begin{bmatrix} X \\ \gamma I_p \end{bmatrix}, \quad \tilde{\mathbf{y}} = \begin{bmatrix} \mathbf{y} \\ \mathbf{0} \end{bmatrix}$$

   Write down the least square errors of the regression model $\tilde{\mathbf{y}} = \tilde{\mathbf{X}}\beta + \epsilon$. It is

   $$||\tilde{\mathbf{y}} - \tilde{\mathbf{X}}\beta||_2^2 = \left\| \begin{bmatrix} \mathbf{y} - \mathbf{X}\beta \\ \gamma\beta \end{bmatrix} \right\|_2^2 = ||\mathbf{y} - \mathbf{X}\beta||_2^2 + \gamma^2 ||\beta||_2^2$$

   Now can you write down the lasso objective function for the regression model $\tilde{\mathbf{y}} = \tilde{\mathbf{X}}\beta + \epsilon$?

3. (10 pts) Computation: This problem involves the Boston data set. To load this data set in R, load the library "MASS" and load the data "Boston", as follows.

   ```
   > library(MASS)
   > data(Boston)
   ```

   We will now try to predict per capita crime rate using the other variables in this data set. In other words, per capita crime rate is the response, and the other variables are the predictors.

   (a) For each predictor, fit a simple linear regression model to predict the response. Describe your results. In which of the models is there a statistically significant association between the predictor and the response? Create some plots to back up your assertions. (open question, think of what are the best plots to support your claim.)

(b) Fit a multiple regression model to predict the response using all of the predictors. Describe your results. For which predictors can we reject the null hypothesis $H_0 : \beta_j = 0$?

(c) How do your results from (a) compare to your results from (b)? Create a plot displaying the univariate regression coefficients from (a) on the x-axis, and the multiple regression coefficients from (b) on the y-axis. That is, each predictor is displayed as a single point in the plot. Its coefficient in a simple linear regression model is shown on the x-axis, and its coefficient estimate in the multiple linear regression model is shown on the y-axis.

(d) Is there evidence of non-linear association between any of the predictors and the response? To answer this question, for each predictor X, fit a model of the form

$$Y = \beta_0 + \beta_1 X + \beta_2 X^2 + \beta_3 X^3 + \epsilon.$$

4. (10 pts) Choose one of the following two problems.

Problem a. Ex. 3.16 from the textbook.

Problem b. Ex. 3.17 from the textbook–Linear methods on the spam data.