Problem 1. (Incremental Stagewise Regression) Read Section 3.8 of the textbook Elements of Statistical Learning, 2nd edition.

A. Implement Algorithm 3.4 as a function incrementalStagewise(x, y, eps, plot = TRUE). (The way the algorithm is described requires to centered all the variables first, so that there is no need to consider an intercept.) Suggest a default value for eps if you can, explaining the rationale for that default value. Produce a plot of the sample path if plot = TRUE as in Figure 3.19 in the same book.

B. Apply your function to synthetic data, similar to the one used in lecture (Part 8). Try a few different configurations and report on 1-3 of them. Each time, compare with the LASSO. (As usual, look at multiple repeats to get a better understanding of what happens 'on average'.)

Problem 2. (MLE distribution in logistic regression) Consider a simple situation with only one predictor variable $x \sim \text{Unif}[0,1]$ and $y \in \{0,1\}$ following a logit model

$$\mathbb{P}(y = 1 \mid x) = \frac{1}{1 + \exp(-\beta_0 - \beta_1 x)}$$

For each $n \in \{20, 50, 100, 200, 500, 1000\}$, repeat the following $R = 1000$ times. With your choice of (β_0, β_1), generate a sample of size n from the model described above. Fit the logistic model to the resulting data and record the value of the MLE $(\widehat{\beta}_0, \widehat{\beta}_1)$. With these R realizations of the MLE produce a histogram with about 50 bins and overlay the asymptotic distribution seen in lecture. Offer some brief comments. (At the end, you will have produced and commented on one such plot for each value of n.)