Extra exercise 7.1
Simulation experiment with OLS and ridge regression with different method for choosing tuning parameters

We continue on the same simulation experiments that were used in extra exercises 3.2-3.4, but we use only ridge regression in addition to OLS.

a) Generate *one* training data sample with $N = 20$ observations, and estimate the model by OLS and ridge regression. For ridge regression, choose the method by

1) 10-fold cross validation

2) Cp, which is equivalent to AIC with a fixed estimate of σ^2 based on a low-bias model, Eqs. (7.26) and (7.30) in the book.

3) AIC as in Eq. (7.29) in the book, with σ^2 estimated by $(1/(n-df)\sum_{i=1}^{i=N}(\widehat{y}_i - y_y)^2$ from the current model.

4) BIC, with σ^2 estimated as for AIC.

The effective number of parameters for ridge regression may be found from Eq. (3.50) in the book, but you can add 1 to account for the intercept.

Plot the values of each selection criteria for various tuning parameter values towards the corresponding effective number of parameters.

b) Perform the same simulation experiments as in extra exercises 3.2b-3.4b, but use only OLS and ridge regression with tuning parameters chosen by the methods above.

c) When k-fold cross validation are used with $k < n$, the division of the data set into k subsets is not unique. There is a randomness in how the data set is divided into separate groups, and different groupings may give slightly different results. The effect of this randomness will be smaller if we compute RMSE by k-fold cross validation for several groupings and then take the average over the grouping. In practice, this may be done by putting the dataset into a random order before dividing it into k groups, and we may call it permuted cross validation.

If you have time, extend exercise b) with permuted cross validation with four permutations.