Tutorial 1

1. Prove the solution of ridge regression estimator

$$\hat{\beta}_R = \min_{\beta}\{\sum_{i=1}^{n}(Y_i - X_i^\top \beta)^2 + \lambda \sum_{k=1}^{p} \beta_k^2$$

is

$$\hat{\beta}_R = (\sum_{i=1}^{n} X_i^\top X_i + \lambda I)^{-1} \sum_{i=1}^{n} X_i Y_i$$

or

$$\hat{\beta}_R = (\mathbf{X}^\top \mathbf{X} + \lambda I)^{-1} \mathbf{X}^\top Y.$$

what about

$$\hat{\beta}_R = \min_{\beta}\{\sum_{i=1}^{n}(Y_i - X_i^\top \beta)^2 + \sum_{k=1}^{p} \lambda_k \beta_k^2$$

where $\lambda_k > 0, k = 1, ..., p$

2. In Example 1.1 of lecture notes chapter 1 (part 1), after removing $\mathbf{x}_5$, fit a new linear regression model. Check whether there is other variables that can be removed (using both T statistics and CV method).

3. In Near Infra-red Calibration for Protein, Fearn (1983) **(dataA)**, use variable selection method to select the model, then apply ridge regression estimate the model. Check the estimated model's prediction error for **(dataB)**