# 4

# Ridge and Lasso regression

Download the LA ozone dataset from the book homepage.

> **Question 4.1.** *Divide the dataset into two groups at random. One group, which we call the training data, containing 2/3 of the observations and one group, which we call the test data, with 1/3 of the observations.*

In the following you are asked to regress the *cube root* of the ozone concentration on the other variables. You should in the following *only* use the training data for the estimation.

> **Question 4.2.** *Compute the best model for each dimension and compute then the test error on the training data.*

> **Question 4.3.** *Compute the ridge and lasso paths. Compute the test error on the test data and plot it as a function of the penalization.*

> **Question 4.4.** *Compute the ridge estimate (for a single $\lambda$) using ordinary least squares by using* `lm`.