

Consider the data frame Auto from library ISLR. It is of interest to predict the car's mileage (mpg) using predictor horsepower. Use cross validation to find the best polynomial model. Use function `cv.glm` from library `boot`. Use the following methods

多項式模型

- Validation set approach with roughly 50% of available data as training set (use `set.seed(1)`).
- Leave-One-Out cross validation.
- 10-fold cross validation

library(TSLL)

validation approach.

`x_train = sample(x=1:10, size=6)` p182
`set.seed(1)`

`trainy = sample(1:92, size=46)`

`m1 = lm(mpg ~ horsepower, Auto, subset = train)`

attach(Auto)

`res = (mpg - predict(m1, Auto))[-train]^2`

`model = mean(res)`

`m2 = lm(mpg ~ poly(horsepower, 2), Auto, subset = train)`

Leave one out cross validation

library(boot)

`glm1 = glm(mpg ~ horsepower, data = Auto)`

`cverr = cv.glm(Auto, glm1)`

`cverr$delta` # MSE or CV from 100 CV