STAT361 Laboratory for Advanced R for Data Science

Lab 10

Sidi Wu

March 30/31, 2023

Department of Statistics and Actuarial Science
Simon Fraser University

MARS: Prediction

Objective

- Refer predict.lm( ) function
- Implement predict method for “mars” objects
- Compare the output of your mars() with the following test file,
  - testpredict.RData (Obtain from the STAT360 repository)

predict.mars()

Your task is to write function make_B()

```
predict.mars <- function(object,newdata) {
  if(missing(newdata) || is.null(newdata)) {
    B <- as.matrix(object$B)
  }
  else {
    tt <- terms(object$formula,data=newdata)
    tt <- delete.response(tt)
    mf <- model.frame(tt,newdata)
    mt <- attr(mf, "terms")
    X <- model.matrix(mt, mf)[,-1] # remove intercept
    B <- make_B(X,object$Bfuncs)
  }
  beta <- object$coefficients
  drop(B %*% beta)
}
```

Object: a mars object

newdata == NULL

- prediction will be done on the training data
- not required construct basis functions
- object$B can be directly used for prediction on training data

newdata != NULL

- prediction will be done on test data
- required to construct basis functions using the Bfuncs list

make_B()

make_B() aims to create set of basis functions with

- New data on covariates
- Information provided in Bfuncs list
- You will be required to use your hinge function (h)
- Refer to the construction of basis functions as a product of hinge functions in lab 5

Test script: testpredict.R

Create a R script called ‘testpredict.R’ and add it to your ‘mars/tests/testthat’ directory,

- It should load ‘testpredict.Rdata’ and call ‘predict.mars( )’ function with following inputs
  - **predict.mars(testmars):**
    - predictions for the same data used to fit the model. Compare the output of this call to the expected output testpredict
  - **predict(testmars,newdata=marstestdata):**
    - predictions from the fitted model in testmars on “new data” in marstestdata
    - same data used to fit the model, but by passing it in as newdata we test that part of the code
    - Compare the output of this call to the expected output testpredict

Next Week

- Labs on next Thursday: office-hour style, no new materials
- No labs or office hour on next Friday (Good Friday holiday)