Ngoc Quang Anh Nguyen – 301355782

Dang Xuan Anh Tran - 301354210

STAT 360 - Advanced R for Data Science

16th April 2021

Multivariate Adaptive Regression Splines

1. Description:

MARS stand for multivariate adaptive regression splines and this function is used as a new method for flexible regression modeling of high dimensional data with the power of continuous models with continuous derivatives.

2. Usage:

source("mars.R")

mars(formula, data, control = mars.control)

3. Arguments:

formula	An object of class "formula" which is a description of the model to be used. Formula interface to specify response and explanatory variables
data	The data frame containing variables in the model. data argument for input data

control	A mars.control object that is a list comprised of elements Mmax, d and trace. Mmax is for the forward, d is for the backward and trace is to print details of the fitting process.

4. Details:

MARS algorithm implements the forward stepwise by incorporating the modifications to recursive partitioning algorithm. The MARS algorithm produces $M_{max}\mathrm{q} = 1$ tensor product (truncated power) spline basis functions that are a subset of the complete tensor product basis with knots located at all distinct marginal data values. Along side with the forward stepwise, the MARS algorithm implements a backwards stepwise deletion strategy, to produce a final set of basis functions. Finally, the knot locations are used to derive a piecewise cubic basis, with continuous first derivatives to produce the final continuous derivative model.

5. Value:

mars returns an object of class "mars" inheritance from class "lm"

An object of class "mars" is a list containing at least the following components:

call: mars() function call

formula: User input formula

y: Response variable

B: Basis function

splits: Best split point

x: Explanatory variable

control: MARS control object

6. Authors:

Ngoc Quang Anh Nguyen – 301355782

Dang Xuan Anh Tran – 301354210

7. References:

Multivariate Adaptive Regression Splines

Author(s): Jerome H. Friedman

Source: The Annals of Statistics , Mar. 1991, Vol. 19, No. 1 (Mar. 1991), pp. 1-67

Published by: Institute of Mathematical Statistics

Stable URL: https://www.jstor.org/stable/2241837

8. See Also:

print.mars(): Which is a method for the MARS object, for better format of the object when printing out to terminal.

summary.mars(): This function will generate a detail summary of the MARS object.

anova.mars(): ANOVA method for MARS object. This function will return the ANOVA decomposition for the output of a MARS object.

predict.mars(): A MARS object method for prediction, including confidence and prediction intervals.

plot.mars(): Plot method for MARS object. This function will generate scatter plot for a MARS object (Fitted vs. Residuals).

9. Examples:

```
source("mars.R")

source("summary.R")

source("print.R")

source("plot.R")

source("predict.R")

source("anova.R")

library(MASS)

library(ISLR)

# TEST 1: Wage data

data("Wage")

mc <- mars.control(Mmax=10)

mout <- mars(wage ~ age + education, data=Wage, control=mc)

ff <- fitted(mout)

p1 <- predict(mout)
```

```
p2 <- predict(mout,newdata=data.frame(age=Wage$age,education=Wage$education))
head(cbind(ff,p1,p2)) # columns should be identical
mout # tests print method
summary.mars(mout) #test summary method
anova.mars(mout) # test anova method
plot.mars(mout) # test plot method
# TEST 2: Trees data
data("trees")
credit = mars.control(Mmax = 10)
cout = mars(Girth ~ Height + Volume, data = trees, control = credit)
ff3 = fitted(cout)
predict3 = predict(cout)
predict4 = predict(cout, newdata = data.frame(Height = trees$Height, Volume = trees$Volume))
head(cbind(ff3, predict3, predict4))
cout
summary.mars(cout)
anova.mars(cout)
plot.mars(aout)
```

```
# TEST 3: Auto data
data("Auto")
auto = mars.control(Mmax = 10)
aout = mars(horsepower ~ mpg + cylinders, data = Auto, control = auto)
ff2 = fitted(aout)
predict1 = predict(aout)
predict2 = predict(aout, newdata = data.frame(mpg = Auto$mpg, cylinders = Auto$cylinders))
head(cbind(ff2, predict1, predict2))
aout
summary.mars(aout)
anova.mars(aout)
plot.mars(aout)
```