Package ‘Kpart’

May 18, 2018

Type    Package
Title   Cubic Spline Fitting with Knot Selection
Version 1.2.2
Date    2018-05-16
Author  Eric Golinko
Depends leaps
Maintainer Eric Golinko <egolinko@gmail.com>
Description Cubic spline fitting along with knot selection, includes support for additional variables.
License GPL (>= 2)
Repository CRAN
Date/Publication 2018-05-18 03:42:57 UTC
RoxygenNote 5.0.1
NeedsCompilation no

R topics documented:

  Kpart-package .................................................. 1
  part ............................................................. 2

Index

Kpart-package    Kpart

Description

Cubic spline regression using the absolute maximum deviate to determine potential knots. This version also includes support for additional independent variables to be included in the model.

Details
This package is intended for use with non-linearly associated data. The function `part` selects points for cubic spline knots using an algorithm to find the absolute maximum deviate from the partition mean, then fits a best fitting model by using the best subset method and maximum adjR2. The function returns the values selected as knots in the model. The function `part(d, outcomeVariable, splineTerm, additionalVars = NULL, K)` takes five arguments. K is a positive integer that indicates how many equally spaced partitions the user would like to search.

--- Recent update includes support for additional variables, 2016-07-23. ---

**Author(s)**

Eric Golinko

Maintainer: egolinko@gmail.com

**References**


---

### `part`

Fits a linear model based on spline terms with additional support for other independent variables.

---

### Description

The user will input a data frame, then designate the variable that is the outcome. Then the spline term is selected along with any other independent variables. Finally, a number K partitions is chosen for the algorithm to search for potential cubic spline knots based on the spline term and partition.

**Usage**

```r
part(d, outcomeVariable, splineTerm, additionalVars = NULL, K)
```
Arguments

- **d**: A data frame data set with column names.
- **outcomeVariable**: The variable from 'd' that is the outcome.
- **splineTerm**: The spline term, inherited from 'd'.
- **additionalVars**: A vector of additional variables to be included in the model.
- **K**: The number of evenly spaced partitions to be searched.

Value

- **fits**: The fitted values of the linear model.
- **xhat**: The entire feature matrix.
- **coefs**: The significant coefficients of the model.
- **adjr2**: The adjusted R^2 value.

Author(s)

Eric Golinko

Examples

```r
## for simple spline model
data(LakeHuron)
d <- data.frame(seq(1875, 1972, 1), LakeHuron)
names(d) <- c('date', 'lh')
fit <- part(d = d, outcomeVariable = 'lh', splineTerm = 'date', K = 20)
fit
plot(d$date, d$lh)
lines(d$date, fit$fits, col = 'red')

## multivariate
data(freeny)
freeny$time <- as.numeric(rownames(freeny))
fit <- part(d = freeny, outcomeVariable = 'y',
   splineTerm = 'time', additionalVars = c('market.potential', 'income.level'), K = 2)
fit$coefs
```
Index

*Topic package
  Kpart-package, 1
*Topic regression
  part, 2

Kpart (Kpart-package), 1
Kpart-package, 1

part, 2