Package ‘STOPES’

June 14, 2019

Type Package
Title Selection Threshold Optimized Empirically via Splitting
Version 0.1
Date 2019-05-15
Author Marinela Capanu, Mihai Giurcanu, Colin Begg, and Mithat Gonen
Maintainer Marinela Capanu <capanum@mskcc.org>
Imports changepoint, glmnet, MASS
Description A variable selection procedure for low to moderate size linear regressions models. This method repeatedly splits the data into two sets, one for estimation and one for validation, to obtain an empirically optimized threshold which is then used to screen for variables to include in the final model.
License GPL-2
NeedsCompilation no
Repository CRAN
Date/Publication 2019-06-14 08:10:03 UTC

R topics documented:

alasso.cv ......................................................... 1
stopes ......................................................... 2

Index

alasso.cv

Description

alasso.cv computes the ALASSO estimator.
stopes

Usage

alasso.cv(x, y)

Arguments

x  n x p covariate matrix
y  n x 1 response vector

Value

alasso.cv returns the ALASSO estimate
alasso the ALASSO estimator

References


Examples

p <- 5
n <- 100
beta <- c(2, 1, 0.5, rep(0, p - 3))
x <- matrix(nrow = n, ncol = p, rnorm(n * p))
y <- rnorm(n) + crossprod(t(x), beta)
lasso.cv(x, y)

stopes

Selection of Threshold OPtimized Empirically via Splitting (STOPES)

Description

stopes computes the STOPES estimator.

Usage

stopes(x, y, m = 20, prop_split = 0.50, prop_trim = 0.20, q_tail = 0.90)

Arguments

x  n x p covariate matrix
y  n x 1 response vector
m  number of split samples, with default value = 20
prop_split proportion of data used for training samples, default value = 0.50
prop_trim proportion of trimming, default prop_trim = 0.20
q_tail proportion of truncation samples across the split samples, default values = 0.90
Value

stopes returns a list with the STOPE estimates via data splitting using 0.25 method and the PELT method:

- `beta_stopes` the STOPE estimate via data splitting
- `J_stopes` the set of active predictors corresponding to STOPEs via data splitting
- `final_cutpoints` the final cutpoint for STOPEs
- `beta_pelt` the STOPE estimate via PELT
- `J_pelt` the set of active predictors corresponding to STOPEs via PELT
- `final_cutpoints_pelt` the final cutpoint for PELT
- `quan_NA` test if the vector of trimmed cutpoints has length 0, with 1 if TRUE and 0 otherwise

Author(s)

Marinela Capanu, Mihai Giurcanu, Colin Begg, and Mithat Gonen

Examples

```r
p <- 5
n <- 100
beta <- c(2, 1, 0.5, rep(0, p - 3))
x <- matrix(nrow = n, ncol = p, rnorm(n * p))
y <- rnorm(n) + crossprod(t(x), beta)
stopes(x, y)
```
Index

alasso.cv, 1
stopes, 2