Package ‘phd’
July 2, 2019

Type Package
Title Permutation Testing in High-Dimensional Linear Models
Version 0.1
Date 2019-06-28
Author Jesse Hemerik, Livio Finos
Maintainer Jesse Hemerik <jesse.hemerik@medisin.uio.no>
Description Provides permutation methods for testing in high-dimensional linear models. The tests are often robust against heteroscedasticity and non-normality and usually perform well under anti-sparsity. See Hemerik and Goeman (2018) <doi:10.1007/s11749-017-0571-1>.
License GNU General Public License
LazyData TRUE
Imports methods, stats, glmnet
NeedsCompilation no
Repository CRAN
Date/Publication 2019-07-02 15:10:03 UTC

R topics documented:

doubleres Permutation test based on double residualization

Index

doubleres

Description

Provides a class of tests for testing in high-dimensional linear models. The tests are robust against heteroscedasticity and non-normality. They often provide good type I error control even under anti-sparsity.
Usage

doubleres(y, X, X1, nperm=2E4, lambda="lambda.min", flip="FALSE", nfolds=10)

Arguments

y
The values of the outcome.

X
The design matrix. If the covariate of interest is included in \( X \), it should be included in the first column. If it is not included in \( X \), then specify \( X1 \). The data do not need to be standardized, since this is automatically done by this function. Do not include a columns of 1’s.

X1
n-vector with the (1-dimensional) covariate of interest. \( X1 \) should only be specified if the covariate of interest is not already included in \( X \).

nperm
The number of random permutations (or sign-flipping maps) used by the test

lambda
The penalty used in the ridge regressions. Default is "lambda.min", which means that the penalty is obtained using cross-validation. One can also enter "lambda.1se", which is an upward-conservative estimate of the optimal lambda.

flip
Default is "FALSE", which means that permutation is used. If "TRUE", then sign-flipping is used.

nfolds
The number of folds used in the cross-validation (in case lambda is determined using cross-validation).

Value

A two-sided p-value.

Examples

set.seed(5193)
n=30

\[ X \leftarrow \text{matrix}(n=n, n=60, \text{rnorm}(n=60)) \]
\[ y \leftarrow X[,1]+X[,2]+X[,3]+\text{rnorm}(n, \text{mean}=0) \quad \#H_0: \text{first coefficient}=0. \text{So H}_0 \text{ is false} \]

doubleres(y, X, nperm=2000, lambda=100, flip="FALSE")

---

**FLhd**

_Freedman-Lane HD_

Description

Provides a class of tests for testing in high-dimensional linear models. The tests are robust against heteroscedasticity and non-normality. They often provide good type I error control even under anti-sparsity.
Usage

\[
\text{FLhd}(y, X, X1, \text{nperm}=2E4, \text{lambda}="\text{lambda.min}"\text{,}\text{flip}="\text{FALSE}"\text{,}\text{nfolds}=10, \text{statistic}="\text{partialcor}"
)\]

Arguments

- \textit{y}
  - The values of the outcome.
- \textit{X}
  - The design matrix. If the covariate of interest is included in \textit{X}, it should be included in the first column. If it is not included in \textit{X}, then specify \textit{X1}. The data do not need to be standardized, since this is automatically done by this function. Do not include a columns of 1's.
- \textit{X1}
  - \text{n-vector with the (1-dimensional) covariate of interest. \textit{X1} should only be specified if the covariate of interest is not already included in \textit{X}.
- \textit{nperm}
  - The number of random permutations (or sign-flipping maps) used by the test
- \textit{lambda}
  - The penalty used in the ridge regressions. Default is "\text{lambda.min}"\text{, which means that the penalty is obtained using cross-validation. One can also enter "\text{lambda.1se}"\text{, which is an upward-conservative estimate of the optimal lambda."
- \textit{flip}
  - Default is "\text{FALSE}"\text{, which means that permutation is used. If "\text{TRUE}"\text{, then sign-flipping is used."
- \textit{statistic}
  - The type of statistic that is used within the permutation test. Either the partial correlation ("\text{partialcor}"\text{ or the semi-partia correlation ("\text{semipartialcor}").
- \textit{nfolds}
  - The number of folds used in the cross-validation (in case \text{lambda} is determined using cross-validation).

Value

A two-sided p-value.

Examples

\begin{verbatim}
set.seed(5193)
n=30

X <- matrix(nr=n,nc=60,rnorm(n*60))
y <- X[,1]+X[,2]+X[,3] + rnorm(n,mean=0)  #H0: first coefficient=0. So H0 is false

FLhd(y, X, nperm=2000, lambda=100, flip="FALSE", statistic="partialcor")
\end{verbatim}
Index

doubleres, 1
FLhd, 2