Package ‘hiertest’

May 5, 2015

Title Convex Hierarchical Testing of Interactions

Version 1.1

Author Jacob Bien, Noah Simon, and Rob Tibshirani

Maintainer Jacob Bien <`jbien@cornell.edu`>

Description Implementation of the convex hierarchical testing (CHT) procedure introduced in Bien, Simon, and Tibshirani (2015) Convex Hierarchical Testing of Interactions. Annals of Applied Statistics. Vol. 9, No. 1, 27-42.

Depends R (>= 3.0.2)

License GPL-3

LazyData true

NeedsCompilation no

Repository CRAN

Date/Publication 2015-05-05 07:53:06

R topics documented:

`hiertest-package` *Convex Hierarchical Testing of Interactions.*

Description

Implementation of the convex hierarchical testing (CHT) procedure introduced in Bien, Simon, and Tibshirani (2015) Convex Hierarchical Testing of Interactions. Annals of Applied Statistics. Vol. 9, No. 1, 27-42. Its main functions are `hiertest` and `estimate.fdr`.

Author(s)

Jacob Bien, Noah Simon, and Rob Tibshirani

References

Bien, Simon, and Tibshirani (2015) Convex Hierarchical Testing of Interactions. Annals of Applied Statistics. Vol. 9, No. 1, 27-42.

`compute.contrasts` *Computes contrasts*

Description

Computes the main effect and interaction contrasts, "w" and "z", respectively.

Usage

```
compute.contrasts(x, y, type = c("Fisher", "simple"))
```

Arguments

`x`	n by p design matrix
`y`	binary (0 or 1) vector of length n indicating class
`type`	determines whether Fisher transform should be applied to interaction contrasts. Default: Fisher

Value

Returns w and z.

`estimate.fdr` *Estimate FDR*

Description

Estimates False Discovery Rate (FDR) based on permutation scheme described in CHT paper (reference below).

Usage

```
estimate.fdr(x, y, lamlist, type = c("Fisher", "simple"), B = 100)
```

Arguments

`x`	n by p design matrix
`y`	binary (0 or 1) vector of length n indicating class
`lamlist`	a vector of cutoffs for the statistics
`type`	determines whether Fisher transform should be applied to interaction contrasts. Default: Fisher. See hiertest for more information.
`B`	number of permutations

Value

A `estfdr` object, which consists of

`ncalled`**:** number of interactions called significant at each cutoff. Set to NA if 0.

`null.ncalled`**:** total number, across all permutations, of (null) interactions rejected at each cutoff

`fdr`**:** estimate of fdr for each cutoff in lamlist. Set to NA if no interactions are rejected at this cutoff

References

Bien, Simon, and Tibshirani (2015) Convex Hierarchical Testing of Interactions. Annals of Applied Statistics. Vol. 9, No. 1, 27-42.

See Also

`hiertest`

Examples

```
# generate some data accoring to the backward model:
set.seed(1)
n <- 200
p <- 50
y <- rep(0:1, each=n/2)
x <- matrix(rnorm(n*p), n, p)
colnames(x) <- c(letters,LETTERS)[1:p]
```

```
# make some interactions between several pairs of variables:
R <- matrix(0.3, 5, 5)
diag(R) <- 1
x[y==1, 1:5] <- x[y==1, 1:5] %*% R
# and a main effect for variables 1 and 3:
x[y==1, 1:5] <- x[y==1, 1:5] + 0.5
testobj <- hiertest(x=x, y=y, type="Fisher")
# look at test statistics
print(testobj)
plot(testobj)
## Not run:
lamlist <- seq(5, 2, length=100)
estfdr <- estimate.fdr(x, y, lamlist, type="Fisher", B=200)
plot(estfdr)
print(estfdr)
# the cutoff lamlist[70] is estimated to have roughly 10% FDR:
estfdr$fdr[70]
# this allows us to reject this many interactions:
nrejected <- estfdr$ncalled[70]
# These are the interactions rejected:
interactions.above(testobj, lamlist[70])

## End(Not run)
```

`hiertest` *Convex Hierarchical Testing Method*

Description

This is the main function, implementing the Convex Hierarchical Testing (CHT) procedure. The CHT procedure produces a set of test statistics for both main effects and interactions with the property that an interaction's statistic is never larger than at least one of its two main effects. This is accomplished by formulating a convex optimization problem that enforces a hierarchical sparsity relationship between the main effects and interactions. The result is that interactions with large main effects receive a "boost" relative to those that do not.

Usage

```
hiertest(x, y, type = c("Fisher", "simple"))
```

Arguments

`x`	n by p design matrix
`y`	binary (0 or 1) vector of length n indicating class
`type`	determines whether Fisher transform should be applied to interaction contrasts. See below for explanation. Default is Fisher and is the recommended choice.

Details

The Convex Hierarchical Testing test statistics are the knots of the CHT optimization problem. That is, the statistic for a given main effect or interaction is the value of lambda at which the corresponding parameter becomes nonzero in the regularization path. Theorem 1 of the CHT paper gives the closed form expression used to compute these knots (recall that for the interaction test statistics, one takes the maximum of the two corresponding knots).

In Section 2.1 of the CHT paper, the raw main effect and interaction contrasts are defined. These are referred to as "w" and "z" in the paper. The main effect contrast "w" is the standard two-sample t-statistic. The interaction contrast "z" is the normalized difference of the Fisher transformed sample correlations between the two classes. If one instead uses `type="simple"`, we simply take for "z" a two-sample statistic on the products of features. We recommend that `type="Fisher"` be used instead of `"simple"`.

Value

A hiertest object, which consists of an ordered list of the main effects and interactions and a vector indicating which of these are interactions.

References

Bien, Simon, and Tibshirani (2015) Convex Hierarchical Testing of Interactions. Annals of Applied Statistics. Vol. 9, No. 1, 27-42.

See Also

`estimate.fdr`

Examples

```
# generate some data accoring to the backward model:
set.seed(1)
n <- 200
p <- 50
y <- rep(0:1, each=n/2)
x <- matrix(rnorm(n*p), n, p)
colnames(x) <- c(letters,LETTERS)[1:p]
# make some interactions between several pairs of variables:
R <- matrix(0.3, 5, 5)
diag(R) <- 1
x[y==1, 1:5] <- x[y==1, 1:5] %*% R
# and a main effect for variables 1 and 3:
x[y==1, 1:5] <- x[y==1, 1:5] + 0.5
testobj <- hiertest(x=x, y=y, type="Fisher")
# look at test statistics
print(testobj)
plot(testobj)
## Not run:
lamlist <- seq(5, 2, length=100)
estfdr <- estimate.fdr(x, y, lamlist, type="Fisher", B=200)
plot(estfdr)
```

```
print(estfdr)
# the cutoff lamlist[70] is estimated to have roughly 10% FDR:
estfdr$fdr[70]
# this allows us to reject this many interactions:
nrejected <- estfdr$ncalled[70]
# These are the interactions rejected:
interactions.above(testobj, lamlist[70])

## End(Not run)
```

`interactions.above` *Get significant interactions above a threshold*

Description

Returns all interactions whose statistics are above a certain level

Usage

```
interactions.above(testobj, lambda)
```

Arguments

`testobj`	output of hiertest
`lambda`	a threshold above which we're interested in interactions

`plot.estfdr` *Plot the estimated FDR*

Description

Plot the estimated FDR

Usage

```
## S3 method for class 'estfdr'
plot(x, ...)
```

Arguments

`x`	output of estimate.fdr
`...`	additional arguments to pass to plot.

`plot.hiertest` *Plot Convex Hierarchical Testing Statistics*

Description

This plots the main effect and interaction statistics from the CHT procedure.

Usage

```
## S3 method for class 'hiertest'
plot(x, ...)
```

Arguments

`x`	output of hiertest.
`...`	additional arguments to `plot`

`print.estfdr` *Print the estimated FDR*

Description

Print the estimated FDR

Usage

```
## S3 method for class 'estfdr'
print(x, ...)
```

Arguments

`x`	output of estimate.fdr
`...`	additional arguments to pass to print.data.frame.

`print.hiertest` *Print Convex Hierarchical Testing Statistics*

Description

This prints the top main effect and interaction statistics from the CHT procedure. Use estimate.fdr to get an appropriate cutoff.

Usage

```
## S3 method for class 'hiertest'
print(x, nshow = 20, ...)
```

Arguments

`x`	output of hiertest
`nshow`	show the largest `nshow` statistics
`...`	additional arguments to `print.data.frame`

Index