Package 'BHPathway'

August 28, 2018

Type Package

Title Bayesian Hierarchical Pathway analysis

Version 1.0

Author Lei Zhang, Swati Biswas

Maintainer Lei Zhang <lxz096120@utdallas.edu>

Description

This function takes a dataset of genotypes, disease phenotypes (case/control), and pathway-gene-SNP information. BHPathway is used to conduct pathway analysis.

License GPL-3

Encoding UTF-8

LazyData true

RoxygenNote 6.1.0.9000

NeedsCompilation no

R topics documented:

`BHPathway` *BHPathway Function*

Description

BHPathway is used to conduct pathway analysis using a combination of Iteratively Weighted Least Squares and Expectation-Maximization algorithms. It outputs p-values at three levels, namely, pathway, genes and SNPs. This function can optionally calculate effective number of parameters for each group of pathways, group of genes, and group of SNPs, and conduct inference based on the effective number of parameters.

Usage

```
BHPathway(type_of_genotype = "numeric", geno, response, pw_gene_snp_info,
  max.iter = 200, FDR.target = 0.1, hierarchical.inference = "FALSE",
  delta.star = 1, use.eff = "FALSE")
```

Arguments

`type_of_genotype`
: can be either "allelic" such as “A G” , or "numeric", e.g., 0, 1, or 2.

`geno`
: SNP genotype matrix with one row per individual and columns listing SNPs in order. If type_of_genotype="allelic", geno has number of columns as 2*number of SNPs. If type_of_genotype="numeric", number of columns in geno is same as the number of SNPs.

`response`
: takes values of 0 or 1.

`pw_gene_snp_info`
: takes three columns, column 1 with pathway ids, column 2 with gene ids, column 3 with SNP ids. The number of rows needs to be same as the total number of SNPs.

`max.iter`
: maximum number of iterations allowed until convergence.

`FDR.target`
: user specified value of FDR boundary for conducting inference, default value is 0.1.

`hierarchical.inference`
: whether inference should be carried out hierarchically at pathway, gene, and SNP levels. Default value is "FALSE" in which case non-hierarchical inference is carried out at pathway level only.

`delta.star`
: used for calculating hierarchical FDR boundary, default value is 1. Needed only when hierarchical.inference=TRUE.

`use.eff`
: this is relevant only when hierarchical.inference="TRUE". If use.eff=="TRUE", cutoffs used in the hierarchical inference in BH (Benjamini-Hochberg) procedure will be based on the effective number of parameters. If use.eff=="FALSE", hierarchical inference uses usual BH thresholds. Default value is "FALSE".

Value

"results" is returned as a list containing following items:

`pw.pvalue`
: vector of p-values for all pathways under study.

`gene.pvalue`
: vector of p-values for all genes under study.

`snp.pvalue`
: vector of p-values for all SNPs under study.

`pw.eff.num`
: effective number of parameters for the group of pathways.

`gene.eff.num`
: effective number(s) of parameters for the groups of genes in each pathway.

`snp.eff.num`
: effective number(s) of parameters for the groups of SNPs in each gene.

`pw.reject.result`
: rejection results at pathway level.

`gene.reject.result`
: rejection results at gene level. This is relevant only when hierarchical.inference="TRUE". "Rejected" means the hypothesis is tested and rejected. "Not tested" means this gene is under a pathway which is not significant, and thus, this gene was not tested.

`snp.reject.result`
: rejection results at SNP level. This is relevant only when hierarchical.inference="TRUE". "Not tested" means this SNP is under a gene or a pathway which is not significant, and thus, this SNP was not tested.

`chosen.qvalue`
: the q-value such that FDR boundary is equal to the user specified value in variable "FDR.target".

References

Zhang, L (2018) A Bayesian Hierarchical Framework for Pathway Analysis in Genome-Wide Association Studies. Ph.D. Dissertation. The University of Texas at Dallas.

Examples

```
## Load example data sets
data(pw_gene_snp_info_eg) ## This consists of information of pathway, gene, SNP mapping information.
data(response_eg)  ## This consists of affection status.
data(geno_eg) ## This consists of SNP genotype matrix.

## Non-Hierarchical inference.
pw.level.without.eff<-BHPathway(type_of_genotype="numeric",geno=geno_eg,response=response_eg,
                                pw_gene_snp_info=pw_gene_snp_info_eg,max.iter=200,
                      FDR.target=0.25,hierarchical.inference="FALSE",delta.star=1,use.eff="FALSE")

## Hierarchical inference without effective number of parameters.
all.level.without.eff<-BHPathway(type_of_genotype="numeric",geno=geno_eg,response=response_eg,
                                 pw_gene_snp_info=pw_gene_snp_info_eg,max.iter=200,
                       FDR.target=0.25,hierarchical.inference="TRUE",delta.star=1,use.eff="FALSE")

## Hierarchical inference with effective number of parameters.
all.level.with.eff<-BHPathway(type_of_genotype="numeric",geno=geno_eg,response=response_eg,
                              pw_gene_snp_info=pw_gene_snp_info_eg,max.iter=200,
                     FDR.target=0.25,hierarchical.inference="TRUE",delta.star=1,use.eff="TRUE")
```

Index