Package 'PAGWAS'

October 12, 2022

Type Package

Title Pathway Analysis Methods for Genomewide Association Data

Version 2.0

Date 2015-12-07

Author Marina Evangelou

Maintainer Marina Evangelou <m.evangelou@ic.ac.uk>

Description Bayesian hierarchical methods for pathway analysis of genomewide association data: Normal/Bayes factors and Sparse Normal/Adaptive lasso. The Frequentist Fisher's product method is included as well.

License GPL (>= 2)

Depends lars, foreach, mnormt

NeedsCompilation no

Repository CRAN

Date/Publication 2015-12-08 14:53:45

R topics documented:

PAGWAS-package *Pathway Analysis Methods for Genomewide Association Data*

Description

Bayesian hierarchical methods for pathway analysis of genomewide association data: Normal/Bayes factors and Sparse Normal/Adaptive lasso. The Frequentist Fisher's method is included as well.

Details

Package:	PAGWAS
Type:	Package
Version:	2.0
Date:	2015-12-02
License:	GPL (>=2)
LazyLoad:	yes

Author(s)

Marina Evangelou Maintainer: Marina Evangelou <m.evangelou@ic.ac.ukt>

References

Evangelou, M., Dudbridge, F., Wernisch, L. (2014). Two novel pathway analysis methods based on a hierarchical model. Bioinformatics 30(5), 690 - 697

Evangelou, M., Rendon, A., Ouhewand, W. H., Wernisch, L., Dudbridge, F. (2012) Comparison of methods for competitive tests of pathway analysis. Plos One 7(7): e41018

create.pathway.df *Creates a pathway data frame*

Description

Returns a data frame with L rows and M columns. L is the number of SNPs in the genotypes data frame and M is the number of tested pathways.

Usage

```
create.pathway.df(genotypes,snps.paths)
```

Arguments

genotypes	Genotype matrix, with L SNPs (columns) and N individuals (rows)
snps.paths	A list with entries the SNP members of each pathway. The size of the list is M

Value

A data frame with columns equal to the number of pathways in the pathway.snps list and rows equal to the number of tested SNPs

See Also

SNPs, genes, snps.to.pathways snps.to.genes

Examples

```
data(SNPs)
data(genes)
data(pathways)
data(genotypes)
snps.genes <- snps.to.genes(snp.info=SNPs,gene.info=genes, distance=0)
pathway.snps <- snps.to.pathways(pathways,snps.genes)
P <- create.pathway.df(genotypes=genotypes,snps.paths=pathway.snps)
```

FM.chi.pvalue *Calculates the Fisher's method p-value for each tested pathway*

Description

Calculates the Fisher's method *p*-value for a set of *p*-values. It returns both the *p*-value and the test statistic value of the Fisher's product method.

Usage

```
FM.chi.pvalue(x)
```

Arguments

x	A vector of *p*-values. These *p*-values can be either gene or SNP *p*-values of a tested pathway

Value

FMstatistic	Fisher's product method test statistic
FMpvalue	Fisher's method *p*-value, computed using the exact distribution of the Fisher's method test statistic which is a *Chi^2* distribution with degrees of freedom twice the size of vector x

References

Evangelou M, Rendon A, Ouwehand WH, Wernisch L, Dudbridge F (2012) Comparison of Methods for Competitive Tests of Pathway Analysis. PLoS ONE 7(7): e41018. doi:10.1371/journal.pone.0041018

See Also

pathways, snps.to.pathways

Examples

```
FM.chi.pvalue(x=c(0.05,0.1))
```

---

genes *A data frame of 20 artificial genes with their chromosomes and positions on the genome*

Description

A data frame with 20 rows and 4 columns.

Usage

```
data(genes)
```

Format

Column names:

`Name` Name of gene

`Start` Start position of gene on the genome

`End` End position of gene on the genome

`Chr` Chromosome of gene

See Also

SNPs

Examples

```
data(genes)
print(genes[1:5,])
```

genotypes *Genotypes for 100 SNPs and 75 individuals*

Description

A data frame with 75 rows (individuals) and 100 columns (SNPs). The entries of the genotype matrix are 0, 1 and 2. There are no missing values.

Usage

```
data(genotypes)
```

See Also

SNPs

Examples

```
data(genotypes)
```

NBF *Normal/Bayes factors method for finding associated pathways*

Description

A vector of the computed Bayes factors for the tested pathways.

Usage

```
NBF(y, G, P, a, b, s2, nu)
```

Arguments

y	Response vector of length N
G	Genotype matrix, with N rows and L columns (number of tested SNPs)
P	Pathway matrix, with L columns and M columns (number of tested pathways)
a	Hyper-parameter of the variance assumed for the integrated out SNP effects
b	Hyper-parameter of the variance assumed for the pathway effects
s2	Hyper-parameter of the Inverse-Chi-squared distribution assumed for the variance of the response vector
nu	Hyper-parameter of the Inverse-Chi-squared distribution assumed for the variance of the response vector

Value

A vector of the computed Bayes factors of the same length as the number of tested pathways

References

Evangelou, M., Dudbridge, F., Wernisch, L. (2014). Two novel pathway analysis methods based on a hierarchical model. Bioinformatics, 30(5), 690 - 697.

Examples

```
## Not run:
data(genotypes)
G=genotypes
data(pathways)
data(SNPs)
data(genes)
snps.genes=snps.to.genes(SNPs,genes,distance=0)
snps.paths=snps.to.pathways(pathways,snps.genes)
P=create.pathway.df(G,snps.paths)
y=rnorm(nrow(G),mean=0,sd=10)
NBF(y,G,P,a,b,s2,nu)
## End(Not run)
```

---

pathways *A list of 2 pathways with their gene members*

Description

A list of two pathways. The gene members of each pathway are given.

Usage

```
data(pathways)
```

See Also

genes, SNPs

Examples

```
data(pathways)
```

SNAL *Sparse Normal/Adaptive lasso method for finding associated pathways*

Description

Sparse Normal/Adaptive lasso method applied for finding the associated pathways. The iterative algorithm suggested by Wipf and Nagarajan (2008) is applied. A vector equal to the number of tested pathways is returned, the zero entries of the vector correspond to the pathways that are not associated. The posterior estimates of the beta coefficients are also returned as they are described by Wipf and Nagarajan (2008).

Usage

```
SNAL(y, G, P, a, s2)
```

Arguments

y	Response vector of length N
G	Genotype matrix, with N rows and L columns (number of tested SNPs)
P	Pathway matrix, with L columns and M columns (number of tested pathways)
a	Hyper-parameter of the variance assumed for the integrated out SNP effects
s2	Variance assumed for the response variable, the tuning parameter of adaptive lasso

Value

gamma.star	Estimates of gamma hyper-parameters
ARD	Posterior estimates of beta coefficients

References

Evangelou, M., Dudbridge, F., Wernisch, L. (2014). Two novel pathway analysis methods based on a hierarchical model. Bioinformatics, 30(5), 690 - 697.

Wipf, D. and Nagarajan, S. (2008). A new view of automatic relevance determination. Advances in Neural Information Processing Systems, 20

See Also

SNAL.calculation

Examples

```
## Not run:
data(genotypes)
G=genotypes
data(pathways)
data(SNPs)
data(genes)
snps.genes=snps.to.genes(SNPs,genes,distance=0)
snps.paths=snps.to.pathways(pathways,snps.genes)
P=create.pathway.df(G,snps.paths)
y=rnorm(nrow(G),mean=0,sd=10)
SNAL(y,G,P,a,s2)
## End(Not run)
```

`SNAL.calculation` *Sparse Normal/Adaptive lasso method for finding associated variables. The SNAL method is applied to the linear regression Y= Phi beta + epsilon*

Description

For more details please read SNAL.

Usage

```
SNAL.calculation(Y, Phi, s2)
```

Arguments

`Y`	Response vector of length N
`Phi`	Design matrix, with N rows and M columns (number of tested variables)
`s2`	Variance assumed for the response variable, the tuning parameter of the adaptive lasso problem

Value

`gamma.star`	Estimates of gamma hyper-parameters
`ARD`	Posterior estimates of beta coefficients

References

Evangelou, M., Dudbridge, F., Wernisch, L. (2014). Two novel pathway analysis methods based on a hierarchical model. Bioinformatics, 30(5), 690 - 697

Wipf, D. and Nagarajan, S. (2008). A new view of automatic relevance determination. Advances in Neural Information Processing Systems, 20

See Also

SNAL

Examples

```
## Not run: SNAL.calculation(Y,Phi,s2=0.5)
```

SNPs *A data frame of 100 artificial SNPs with their chromosomes and positions on the genome*

Description

A data frame with 100 rows and 3 columns.

Usage

```
data(SNPs)
```

Format

Column names:

`Name` SNP name

`Position` Position of SNP on the genome

`Chr` Chromosome of the SNP

See Also

genes, genotypes

Examples

```
data(SNPs)
print(SNPs[1:5,])
```

`snps.to.genes` *Assigns SNPs to genes*

Description

Assigns SNPs to genes based on their physical distance.

Usage

```
snps.to.genes(snp.info, gene.info, distance)
```

Arguments

`snp.info`	A data frame with 3 columns with names: Name, Position and Chr that correspond to the SNP name, its position on the genome and its chromosome, respectively
`gene.info`	A data frame with 4 columns with names: Name, Start, End and Chr that correspond to the gene name, start and end positions on the genome and its chromosome, respectively
`distance`	A number that corresponds to the distance below and above the Start and End positions of the gene that all SNPs in that region should be assigned to the gene

Value

A list of the same size as the number of genes of the gene.info data frame. The names of the SNPs assigned to each gene are returned

See Also

`SNPs`, `genes`, `snps.to.pathways`

Examples

```
data(SNPs)
data(genes)
snps.to.genes(snp.info=SNPs,gene.info=genes,distance=50)
```

snps.to.pathways *Assigns SNPs to pathways*

Description

Assigns SNPs to pathways, using the pathway gene members and the SNPs assigned to each gene.

Usage

```
snps.to.pathways(pathways,gene.snps)
```

Arguments

pathways	A list of pathways with their gene members
gene.snps	A list of genes with the SNPs assigned to them according to their physical distance on the genome

Value

A list of the same size as the number of pathways in the pathway list. The names of the SNPs assigned to each pathway are returned. Empty pathways are also returned.

See Also

SNPs, genes, snps.to.genes

Examples

```
data(SNPs)
data(genes)
data(pathways)
snps.genes <- snps.to.genes(snp.info=SNPs,gene.info=genes, distance=50)
pathway.snps <- snps.to.pathways(pathways,snps.genes)
```

Index