

Package ‘topologyGSA’

July 9, 2015

Version 1.4.5

Date 2015-07-09

Title Gene Set Analysis Exploiting Pathway Topology

Author Sofia Massa, Gabriele Sales

Maintainer Gabriele Sales <gabriele.sales@unipd.it>

Description Using Gaussian graphical models we propose a novel approach to perform pathway analysis using gene expression. Given the structure of a graph (a pathway) we introduce two statistical tests to compare the mean and the concentration matrices between two groups. Specifically, these tests can be performed on the graph and on its connected components (cliques).

Imports graph, gRbase (>= 1.7), qqgraph, fields

Suggests Rgraphviz

License GPL-3

LazyLoad yes

NeedsCompilation no

Repository CRAN

Date/Publication 2015-07-09 18:13:07

R topics documented:

clique.mean.test	2
clique.var.test	3
dag_bcell	5
pathway.mean.test	5
pathway.var.test	7
print	8
render.significant.cliques	9
y1	10
y2	10

Index	12
--------------	-----------

clique.mean.test *Comparison of all the cliques of a pathway.*

Description

Test the equality of the means in two experimental conditions for all the cliques of a pathway, conditioning on the results of the test on the concentration matrices

Usage

```
clique.mean.test(y1,y2,dag,alpha,perm.num=1000,paired=FALSE)
```

Arguments

y1	a matrix with n1 individuals (rows) in the first experimental condition and p genes (columns).
y2	a matrix with n2 individuals (rows) in the second experimental condition and p genes (columns). The genes in the two experimental conditions must be the same.
dag	graphNEL object, a directed acyclic graph (DAG) corresponding to the pathway of interest. See package gRbase for more details.
alpha	significance level of the test.
perm.num	number of permutation used by the test.
paired	logical flag. If TRUE statistical tests for paired samples is calculated and the test on the variances is not performed.

Details

The function tests the equality of the means of each clique of a pathway in two experimental conditions, conditioning on the results of `clique.var.test`. The graph of a pathway is first converted into a DAG, then the moral graph is obtained and if the latter graph is decomposable then the test is performed on all its cliques. If the moral graph is not decomposable, its triangulated version is obtained and then the test is performed on all its cliques.

The expression data may contain some genes differing from those in the pathway: in such case the function automatically takes the intersection between the two gene sets.

This function requires gRBase and qpgraph packages.

Value

p.value	p-value of the test on the mean for each clique.
p.value.var	p-value of the test on the variance for each clique. If paired is TRUE, this is not returned.
cliques	the list of the corresponding cliques.
var.equal	logical, a vector with a TRUE for a significant clique and a FALSE otherwise. If paired is TRUE, this is not returned.

t.value	the observed values of the statistic to test the mean for each clique.
lambda.value	the observed values of the statistic to test the variance for each clique. If paired is TRUE, this is not returned.
graph	the triangularized and moralized graph.

Author(s)

M. Sofia Massa, Gabriele Sales

References

Massa, M.S., Chiogna, M., Romualdi, C. (2010). Gene set analysis exploiting the topology of a pathway. BMC Systems Biology, 4:121 <http://www.biomedcentral.com/1752-0509/4/121>

See Also

[pathway.var.test](#), [pathway.mean.test](#), [clique.var.test](#),

Examples

```
data(examples)
clique.mean.test(y1, y2, dag_bcell, 0.05)
```

clique.var.test *Testing the concentration matrices of all cliques of a pathway.*

Description

Test the equality of concentration matrices in two experimental conditions for all cliques of a pathway

Usage

```
clique.var.test(y1,y2,dag,alpha)
```

Arguments

y1	a matrix with n1 individuals (rows) in the first experimental condition and p genes (columns).
y2	a matrix with n2 individuals (rows) in the second experimental condition and p genes (columns). The genes in the two experimental conditions must be the same.
dag	graphNEL object, a directed acyclic graph (DAG) corresponding to the pathway of interest. See package gRbase for more details.
alpha	significance level of the test.

Details

The function tests the equality of the concentration matrices of each clique of a pathway in two experimental conditions. The graph of a pathway is first converted into a DAG, then the moral graph is obtained and if the latter graph is decomposable then the test is performed on all its cliques. If the moral graph is not decomposable, its triangulated version is obtained and then the test is performed on all its cliques.

The expression data may contain some genes differing from those in the pathway: in such case the function automatically takes the intersection between the two gene sets.

This function requires gRBase and qpgraph packages.

Value

p.value	the observed value of the significance level.
cliques	the list of cliques where the test is performed.
var.equal	logical, a vector with a TRUE for a significant clique and a FALSE otherwise.
lambda.var	the observed value of the statistic.
graph	the triangulated and moral graphs.

Author(s)

M. Sofia Massa, Gabriele Sales

References

Massa, M.S., Chiogna, M., Romualdi, C. (2010). Gene set analysis exploiting the topology of a pathway. BMC Systems Biology, 4:121 <http://www.biomedcentral.com/1752-0509/4/121>

See Also

[pathway.var.test](#), [pathway.mean.test](#), [clique.mean.test](#),

Examples

```
data(examples)

clique.var.test(y1, y2, dag_bcell, 0.05)
```

dag_bcell

*DAG representation of the B cell receptor signaling pathway***Description**

The B cell receptor signaling pathway (BCR) converted to a DAG. In case of nodes with multiple genes, the most differentially expressed between groups of patients defined in y1 (acute lymphocytic leukemia patients with BCR/ABL gene rearrangement) and y2 (acute lymphocytic leukemia patients with BCR/ABL gene rearrangement) matrices.

Chiaretti S., Li X., Gentleman R., Vitale A., Wang K.S., Mandelli F., Fo R., Ritz J. (2005) Gene expression profiles of B-lineage adult acute lymphocytic leukemia reveal genetic patterns that identify lineage derivation and distinct mechanisms of transformation. Clin Cancer Res, 11:7209-7219. <http://clincancerres.aacrjournals.org/content/11/20/7209.long>

Massa, M.S., Chiogna, M., Romualdi, C. (2010). Gene set analysis exploiting the topology of a pathway. BMC Systems Biology, 4:121 <http://www.biomedcentral.com/1752-0509/4/121>

pathway.mean.test

*Testing the means of a pathway.***Description**

Test the equality of means in two experimental conditions for a pathway, conditioning on the result of the test on the concentration matrices

Usage

```
pathway.mean.test(y1,y2,dag,alpha,perm.num=10000,variance=TRUE,paired=FALSE)
```

Arguments

y1	a matrix with n1 individuals (rows) in the first experimental condition and p genes (columns).
y2	a matrix with n2 individuals (rows) in the second experimental condition and p genes (columns). The genes in the two experimental conditions must be the same.
dag	graphNEL object, a directed acyclic graph (DAG) corresponding to the pathway of interest. See package gRbase for more details.
alpha	significance level of the test.
variance	logical flag. If TRUE estimated variances s1 and s2 are reported.
perm.num	number of permutations of the test.
paired	logical flag. If TRUE Hotelling test for paired samples is calculated and the test on the variances is not performed.

Details

The function tests the equality of the means of two experimental conditions, conditioning on the result of `pathway.var.test`.

The expression data may contain some genes differing from those in the pathway: in such case the function automatically takes the intersection between the two gene sets.

This function requires `gRBase` and `qpgraph` packages.

Value

<code>p.value</code>	the p-value of the test on the mean.
<code>p.value.var</code>	the p-value of the test on the variance. If <code>paired</code> is <code>TRUE</code> , this is not returned.
<code>lambda.value</code>	the observed value of the test statistic for the variance. If <code>paired</code> is <code>TRUE</code> , this is not returned.
<code>qchisq.value</code>	the 1 - alpha quantile value of the null distribution of the test statistic on the variance. If <code>paired</code> is <code>TRUE</code> , this is not returned.
<code>cli.moral</code>	a list containing the cliques of the moral graph.
<code>var.equal</code>	logical flag. If <code>TRUE</code> variances are heteroschedastic. If <code>paired</code> is <code>TRUE</code> , this is not returned.
<code>graph</code>	a <code>graphNEL</code> object representing the moral graph.
<code>df.var</code>	the degrees of freedom of the null distribution for the variance test. If <code>paired</code> is <code>TRUE</code> , this is not returned.
<code>df.mean</code>	the degrees of freedom of the null distribution for the mean test. If <code>paired</code> is <code>TRUE</code> , this is not returned.
<code>t.value</code>	the observed value of the test statistic for the mean.

Author(s)

M. Sofia Massa, Gabriele Sales

References

Massa, M.S., Chiogna, M., Romualdi, C. (2010). Gene set analysis exploiting the topology of a pathway. *BMC Systems Biology*, 4:121 <http://www.biomedcentral.com/1752-0509/4/121>

See Also

[pathway.var.test](#), [clique.var.test](#), [clique.mean.test](#),

Examples

```
data(examples)
```

```
pathway.mean.test(y1, y2, dag_bcell, 0.05, 100)
```

pathway.var.test *Testing the concentration matrices of a pathway.*

Description

Test the equality of concentration matrices in two experimental conditions for a pathway

Usage

```
pathway.var.test(y1,y2,dag,alpha,variance=FALSE,s1=NULL,s2=NULL)
```

Arguments

y1	a matrix with n1 individuals (rows) in the first experimental condition and p genes (columns).
y2	a matrix with n2 individuals (rows) in the second experimental condition and p genes (columns). The genes in the two experimental conditions must be the same.
dag	graphNEL object, directed acyclic graph (DAG) corresponding to the pathway of interest. See package gRbase for more details.
alpha	significance level of the test.
variance	logical flag. If TRUE the estimates of the covariance matrices are included in the result.
s1	y1 covariance matrix estimation.
s2	y2 covariance matrix estimation.

Details

The graph of a pathway is first converted into a DAG and then into a moral graph. The data is modelled with two Gaussian graphical models with zero mean and graph provided by the moral graph. The function tests the equality of the two concentration matrices (inverse of the covariance matrices).

The expression data may contain some genes differing from those in the pathway: in such case the function automatically takes the intersection between the two gene sets.

A necessary condition for the existence of the covariance estimates is that the number of statistical units (samples) is greater than the number of variables. If this is not the case, penalized techniques for estimating $\hat{\Sigma}_1^{-1}$ and $\hat{\Sigma}_2^{-1}$ have to be employed, that are currently not provided by the package. In these cases, one can perform penalized estimation of $\hat{\Sigma}_1^{-1}$ and $\hat{\Sigma}_2^{-1}$ outside **topologyGSA**, and then provide such estimates as input arguments to the function `pathway.var.test` to compute the value of the test for homogeneity. In this case, computation of the p-value deserves attention, as standard results on the asymptotic distribution of the test statistic may no longer be valid. Therefore, computation of the p-value has to be dealt with by the user.

This function requires gRBase and qpgraph packages.

Value

<code>lambda.value</code>	the observed value of the test statistic.
<code>qchisq.value</code>	the 1 - alpha quantile value of the null distribution of the test statistic on the variance.
<code>p.value</code>	the observed value of the significance level of the test.
<code>cli.moral</code>	a list containing the cliques of the moral graph.
<code>var.equal</code>	logical flag. If TRUE variances are heteroschedastic.
<code>graph</code>	a graphNEL object representing the moral graph.
<code>df</code>	the degrees of freedom of the null distribution.
<code>s1</code>	if variance=TRUE, the estimate of y1 covariance.
<code>s2</code>	if variance=TRUE, the estimate of y2 covariance.

Author(s)

M. Sofia Massa, Gabriele Sales

References

- Massa, M.S., Chiogna, M., Romualdi, C. (2010). Gene set analysis exploiting the topology of a pathway. *BMC Systems Biology*, 4:121 <http://www.biomedcentral.com/1752-0509/4/121>
- Lauritzen, S.L. (1996). *Graphical models*. Clarendon Press, Oxford.

See Also

[pathway.mean.test](#), [clique.var.test](#), [clique.mean.test](#).

Examples

```
data(examples)

pathway.var.test(y1, y2, dag_bcell, 0.05)
```

print

Print results of statistical tests.

Description

This functions prints a summary of the results of each statistical test.

Usage

```
## S3 method for class 'clique.var.test'  
print(x, ...)  
## S3 method for class 'clique.mean.test'  
print(x, ...)  
## S3 method for class 'pathway.var.test'  
print(x, ...)  
## S3 method for class 'pathway.mean.test'  
print(x, ...)
```

Arguments

x	the result of a statistical test
...	ignored

Author(s)

M. Sofia Massa, Gabriele Sales

render.significant.cliques

Graphical display of the significant cliques

Description

This function highlights the significant cliques on the moralized and triangulated graph. Different colors correspond to different values of $-\log(\text{pvalue})$

Usage

```
render.significant.cliques(info, alpha)
```

Arguments

info	the result of <code>clique.var.test</code> or <code>clique.mean.test</code> .
alpha	p-value cutoff to define significant cliques

References

Massa, M.S., Chiogna, M., Romualdi, C. (2010). Gene set analysis exploiting the topology of a pathway. BMC Systems Biology, 4:121 <http://www.biomedcentral.com/1752-0509/4/121>

See Also

[clique.var.test](#), [clique.mean.test](#).

Examples

```
data(examples)

info <- clique.var.test(y1, y2, dag_bcell, 0.1)
## Not run: render.significant.cliques(info)
```

y1	<i>Gene expression signatures in acute lymphocytic leukemia for BCR signaling pathway</i>
----	---

Description

Several distinct genetic mechanisms lead to acute lymphocytic leukemia (ALL) malignant transformations deriving from distinct lymphoid precursor cells that have been committed to either T-lineage or B-lineage differentiation. Chromosome translocations and molecular rearrangements are common events in B-lineage ALL and reflect distinct mechanisms of transformation. The B cell Receptor (BCR/ABL) gene rearrangement occurs in about 25% of cases in adult ALL, and much less frequently in pediatric ALL. y1 dataset contains gene expression signatures of BCR signaling pathway (composed by 35 gene products, n.columns = 35) in acute lymphocytic leukemia (ALL) cells associated with presence of BCR/ABL gene rearrangement (n. rows = 37). Expression values are normalized according to rma and quantile normalization.

Chiaretti S., Li X., Gentleman R., Vitale A., Wang K.S., Mandelli F., Fo R., Ritz J. (2005) Gene expression profiles of B-lineage adult acute lymphocytic leukemia reveal genetic patterns that identify lineage derivation and distinct mechanisms of transformation. Clin Cancer Res, 11:7209-7219. <http://clincancerres.aacrjournals.org/content/11/20/7209.long>

Massa, M.S., Chiogna, M., Romualdi, C. (2010). Gene set analysis exploiting the topology of a pathway. BMC Systems Biology, 4:121 <http://www.biomedcentral.com/1752-0509/4/121>

y2	<i>Gene expression signatures in acute lymphocytic leukemia for BCR signaling pathway</i>
----	---

Description

Several distinct genetic mechanisms lead to acute lymphocytic leukemia (ALL) malignant transformations deriving from distinct lymphoid precursor cells that have been committed to either T-lineage or B-lineage differentiation. Chromosome translocations and molecular rearrangements are common events in B-lineage ALL and reflect distinct mechanisms of transformation. The B cell Receptor (BCR/ABL) gene rearrangement occurs in about 25% of cases in adult ALL, and much less frequently in pediatric ALL. y2 dataset contains gene expression signatures of BCR signaling pathway (composed by 35 gene products, n.columns = 35) in acute lymphocytic leukemia (ALL) cells not associated with BCR/ABL gene rearrangement (n. rows = 41). Expression values are normalized according to rma and quantile normalization.

Chiaretti S., Li X., Gentleman R., Vitale A., Wang K.S., Mandelli F., Fo R., Ritz J. (2005) Gene expression profiles of B-lineage adult acute lymphocytic leukemia reveal genetic patterns that identify lineage derivation and distinct mechanisms of transformation. *Clin Cancer Res*, 11:7209-7219. <http://clincancerres.aacrjournals.org/content/11/20/7209.long>

Massa, M.S., Chiogna, M., Romualdi, C. (2010). Gene set analysis exploiting the topology of a pathway. *BMC Systems Biology*, 4:121 <http://www.biomedcentral.com/1752-0509/4/121>

Index

`clique.mean.test`, [2](#), [4](#), [6](#), [8](#), [9](#)

`clique.var.test`, [3](#), [3](#), [6](#), [8](#), [9](#)

`dag_bcell`, [5](#)

`pathway.mean.test`, [3](#), [4](#), [5](#), [8](#)

`pathway.var.test`, [3](#), [4](#), [6](#), [7](#)

`print`, [8](#)

`render.significant.cliques`, [9](#)

`y1`, [10](#)

`y2`, [10](#)