

Package ‘POST’

April 12, 2018

Type Package

Title Projection onto Orthogonal Space Testing for High Dimensional Data

Version 1.2.0

Author Xueyuan Cao <xueyuan.cao@stjude.org> and Stanley.pounds <stanley.pounds@stjude.org>

Maintainer Xueyuan Cao <xueyuan.cao@stjude.org>

Description Perform orthogonal projection of high dimensional data of a set, and statistical modeling of phenotype with projected vectors as predictor.

Depends R (>= 3.4.0)

Imports stats, CompQuadForm, Matrix, survival, Biobase, GSEABase

License GPL (>= 2)

biocViews Microarray, GeneExpression

LazyLoad yes

NeedsCompilation no

R topics documented:

POST-package	1
POSTcoxph	2
POSTglm	4
sampExprSet	5
sampGeneSet	5

Index	6
--------------	----------

POST-package	<i>Projection onto Orthogonal Space Testing for High Dimensional Data</i>
--------------	---

Description

a tool to identify gene sets that are associated with a endpoint of interest by projecting the original data of a set to an orthogonal subspace.

Details

```

Package:   POST
Type:     Package
Version:  0.99.3
Date:    2016-11-2
License:  GPL (>=2)
LazyLoad: yes

```

The POST (Projection onto Orthogonal Space Testing) is performed by calling function *POSTglm* for linear model or *POSTcoxph* for Cox proportional hazard model. The genomic data such as gene expression is passed through ExpressionSet; The gene set definition is defined by a gene set collection.

Author(s)

Xueyuan Cao <xueyuan.cao@stjude.org>, Stanley Pounds <stanley.pounds@stjude.org>
 Maintainer: Xueyuan Cao <xueyuan.cao@stjude.org>

References

Hotelling H. (1936). Relations between two sets of variables. *Biometrika*, 28, 321-327

Examples

```

## load data
data(sampExprSet)
data(sampGeneSet)
test<-POSTglm(exprSet=sampExprSet,
              geneSet=sampGeneSet,
              lamda=0.95,
              seed=13,
              nboots=100,
              model='Group ~ ',
              family=binomial(link = "logit"))

```

POSTcoxph

POST Analysis of Cox Proportional Hazard Model

Description

Perform POST analysis with Cox proportional hazard model

Usage

```

POSTcoxph (exprSet, geneSet, lamda = 0.95, nboots = 100,
           model = "Surv(EFSTIME, EFSCENSOR) ~ strata(arm2)",
           seed = 13, ...)

```

Arguments

exprSet	an ExpressionSet class contains minimum of exprs (expression matrix) of genomic data such as gene expression and phenoData (AnnotatedDataFrame of end point data). Please refer to Biobase for details on how to create such an ExpressionSet expression set.
geneSet	a GeneSetCollection class with minimum of setName and geneIDs for each GeneSet. Please refer to <i>GSEABase</i> for how to create such a GeneSetCollection class.
lamda	a predefined fraction of 0 to 1: the fraction of variation retained in the selected orthogonal projections
nboots	number of bootstrap samples, default=100
model	a character string to define model
seed	seed of random number generator
...	control arguments to be used in Cox proportional hazard model, default control arguments is used if they are not supplied directly

Details

The function performs POST analysis for association of gene set with time to event endpoint in Cox proportional hazard model framework.

Value

a data frame with 5 columns

GeneSet	Gene set id
Nprobe	numbe of probes in the ExpressionSet annotated to the gene set
Nproj	number of projected vectors included in the POST analysis
Stat	statistic of quadratic form
p.value	p value of generalized Chi-square approximation

Author(s)

Xueyuan Cao <Xueyuan.cao@stjude.org>, Stanley Pounds <stanley.pounds@stjude.org>

Examples

```
## load data
data(sampExprSet)
data(sampGeneSet)
test<-POSTcoxph(exprSet=sampExprSet,
                geneSet=sampGeneSet,
                lamda=0.95,
                seed=13,
                nboots=100,
                model='Surv(time, censor) ~ ')
```

POSTglm

*POST of Generalized Linear Model***Description**

Perform POST analysis with linear model

Usage

```
POSTglm (exprSet, geneSet, lamda = 0.95, seed = 13, nboots = 100,
        model = "Group ~ ", family = binomial(link = "logit"), ...)
```

Arguments

exprSet	an ExpressionSet class contains minimum of exprs (expression matrix) of genomic data such as gene expression and phenoData (AnnotatedDataFrame of end point data). Please refer to Biobase for details on how to create such an ExpressionSet expression set.
geneSet	a GeneSetCollection class with minimum of setName and geneIDs for each GeneSet. Please refer to <i>GSEABase</i> for how to create such a GeneSetCollection class.
lamda	a predefined fraction of 0 to 1: the fraction of variation retained in the selected orthogonal projections
seed	seed of random number generator
nboots	number of bootstrap samples, default=100
model	a character string to define the model
family	a description of the error distribution and link function to be used in the model. For glm, this can be a character string naming a family function, a family function or the result of a call to a family function.
...	control arguments to be passed to generalized linear model, default control argument are used if they are not supplied directly

Details

The function performs POST analysis for association of gene sets with phenotype in generalized linear model framework.

Value

a data frame with 5 columns

GeneSet	Gene set id
Nprobe	numbe of probes in the ExpressionSet annotated to the gene set
Nproj	number of projected vectors included in the POST analysis
Stat	statistic of quadratic form
p.value	p value of generalized Chi-square approximation

Author(s)

Xueyuan Cao <Xueyuan.cao@stjude.org>, Stanley Pounds <stanley.pounds@stjude.org>

Examples

```
## load data
data(sampExprSet)
data(sampGeneSet)
test<-POSTglm(exprSet=sampExprSet,
              geneSet=sampGeneSet,
              lamda=0.95,
              seed=13,
              nboots=100,
              model='Group ~ ',
              family=binomial(link = "logit"))
```

sampExprSet

Example of Expression Set

Description

an ExpressionSet class contains minimum of exprs (expression matrix) of gene expression and phenoData (AnnotatedDataFrame of end point data).

Usage

```
data(sampExprSet)
```

Value

an simulated ExpressionSet contains conceptual data of 60 expression features for 40 subjects. The phenotype data has 4 columns for the same 40 subjects.

sampGeneSet

Example of Expression Set

Description

a Gene Set Collection contains minimum of GeneSet definition.

Usage

```
data(sampGeneSet)
```

Value

a gene set collection of 4 gene sets with 60 unique probes

Index

*Topic **misc**

sampExprSet, [5](#)

sampGeneSet, [5](#)

*Topic **package**

POST-package, [1](#)

POST-package, [1](#)

POSTcoxph, [2](#)

POSTglm, [4](#)

sampExprSet, [5](#)

sampGeneSet, [5](#)