Package 'ASGSCA'

May 15, 2026

Type Package

Title Association Studies for multiple SNPs and multiple traits using Generalized Structured Equation Models

Version 1.47.0

Date 2014-07-30

Author Hela Romdhani, Stepan Grinek , Heungsun Hwang and Aurelie Labbe.

Maintainer Hela Romdhani <hela.romdhani@mcgill.ca>

Description The package provides tools to model and test the association between multiple genotypes and multiple traits, taking into account the prior biological knowledge. Genes, and clinical pathways are incorporated in the model as latent variables. The method is based on Generalized Structured Component Analysis (GSCA).

License GPL-3

Imports Matrix, MASS

biocViews StructuralEquationModels

Suggests BiocStyle

git_url https://git.bioconductor.org/packages/ASGSCA

git_branch devel

git_last_commit 23cd5a0

git_last_commit_date 2026-04-28

Repository Bioconductor 3.24

Date/Publication 2026-05-15

Contents

ASGSCA-package 2
B0 3
GenPhen 3
GSCA 4
GSCAestim 5
QCAHS 7
ResQCAHS 7
W0 8

Index 9

ASGSCA-package *Association Studies for multiple SNPs and multiple traits using Generalized Structured Component Analysis*

Description

The package provides tools to model and test the association between multiple genotypes and multiple traits, taking into account the prior biological knowledge. Functional genomic regions, e.g., genes, and clinical pathways are incorporated in the model as latent variables that are not directly observed. The method is based on Generalized Structured Component Analysis (GSCA).

Details

Package:	ASGSCA
Type:	Package
Version:	1.0
Date:	2014-07-30
License:	GPL-3

Author(s)

Hela Romdhani, Stepan Grinek, Heungsun Hwang and Aurelie Labbe.

Maintainer: Hela Romdhani <hela.romdhani@mcgill.ca>

References

Romdhani, H., Hwang, H., Paradis, G., Roy-Gagnon, M.-H. and Labbe, A. (2014). Pathway-based Association Study of Multiple Candidate Genes and Multiple Traits Using Structural Equation Models. Submitted.

Examples

```
data(GenPhen)
W0 <- matrix(c(rep(1,2),rep(0,8),rep(1,2),rep(0,8),rep(1,3),rep(0,7),rep(1,2)),nrow=8,ncol=4)
B0 <- matrix(c(rep(0,8),rep(1,2),rep(0,3),1,rep(0,2)),nrow=4,ncol=4)

#Estimation only
GSCA(GenPhen,W0, B0,estim=TRUE,path.test=FALSE)
#Estimation and test for all the path coefficients in the model
GSCA(GenPhen,W0, B0,estim=TRUE,path.test=TRUE)
#Test only
GSCA(GenPhen,W0, B0,estim=FALSE,path.test=TRUE)
#Give names to the latent variables
GSCA(GenPhen,W0, B0,latent.names=c("Gene1","Gene2","Clinical pathway 1","Clinical pathway 2"),
estim=TRUE,path.test=TRUE)
#Testing only a subset of path coefficients
GSCA(GenPhen,W0, B0,estim=FALSE,path.test=TRUE,path=matrix(c(1,2,3,4),ncol=2))
```

B0 — *Matrix indicating connections between the latent variables for QCAHS data.*

Description

Matrix indicating connections between the latent variables in the path model fitted to QCAHS data, used in the vignette to estimate the path coefficients of the model.

Usage

```
data(B0)
```

Format

A square matrix of dimension 28.

GenPhen — *Dataset to test GSCAestim and GSCA functions.*

Description

Simulated data (for 999 individuals) of 4 SNPs mapped to 2 different genes and 4 traits involved in 2 different clinical pathways. The data is simulated such that one of the traits is involved in both clinical pathways and that one gene is connected to one of the clinical pathways and the other to both of them. See Figure 2, scenario (g), in Romdhani et al. (2014) for details.

Usage

```
data(GenPhen)
```

Format

A data frame of 8 columns and 999 rows.

References

Romdhani, H., Hwang, H., Paradis, G., Roy-Gagnon, M.-H. and Labbe, A. (2014). Pathway-based Association Study of Multiple Candidate Genes and Multiple Traits Using Structural Equation Models, submitted.

GSCA *Association test for multiple genotypes and multiple traits using Generalized Structured Component Analysis (GSCA)*

Description

For a specified structural equation model with latent variables relating the traits and the genotypes, the function GSCA gives estimates of the parameters of the model and performs permutation tests for the association between multiple genotypes and multiple traits (see Romdhani et al., 2014).

Usage

```
GSCA(data,W0, B0,latent.names=NULL,estim=TRUE,path.test=TRUE,path=NULL,nperm=1000)
```

Arguments

data	data frame containing the observed variables (genotypes and traits).
W0	matrix with 0's and 1's indicating connections between the observed variables (genotypes and traits) and the latent variables (genes and clinical pathways). The rows correspond to the observed variables in the same order as in data; the columns to the latent variables. A value of 1 indicates an arrow from the observed variable in the row to the latent variable in the column.
B0	square matrix with 0's and 1's indicating connections among latent variables (genes and clinical pathways). Both rows and columns correspond to the latent variables. A value of 1 indicates an arrow directed from the latent variable in the row to the latent variable in the column.
latent.names	optional vector of characters containing names for the latent variables that will be displayed in the results. If NULL c("Latent1","Latent2",...) will be used. Default is NULL.
estim	logical. If TRUE the estimates of the weight and path coefficient are returned. Default is TRUE.
path.test	logical. If TRUE tests for path coefficients are performed. Default is TRUE.
path	an optional matrix with 2 columns indicating particular connections to be tested. Each row contains the indices of the two latent variables (the gene and the clinical pathway) corresponding to the connection to be tested. If NULL, the test is performed for all gene-clinical pathway connections specified in the model. Default is NULL.
nperm	number of permutations. Default is 1000.

Value

If estim is TRUE, returns a list with 2 items:

Weight	Matrix of the same dimension as W0 with 1's replaced by weight coefficients estimates.
Path	Matrix of the same dimension as B0 with 1's replaced by path coefficients estimates.

If, path.test is TRUE and path is NULL, the function returns a matrix of the same dimensions as B0 with 1's replaced by the corresponding p-values and 0's replaced by NA's. If path.test is TRUE and path is not NULL, only p-values for the specified path coefficients are returned.

Author(s)

Hela Romdhani, Stepan Grinek, Heungsun Hwang and Aurelie Labbe.

References

Romdhani, H., Hwang, H., Paradis, G., Roy-Gagnon, M.-H. and Labbe, A. (2014). Pathway-based Association Study of Multiple Candidate Genes and Multiple Traits Using Structural Equation Models. Submitted.

Examples

```
#Scenario (g) in Romdhani et al. (2014): 4 SNPs mapped to 2 genes and 4
#traits involved in 2 clinical pathways
#In total: 8 observed variables and 4 latent variables.
#One of the traits is involved in both clinical pathways.
#One gene is connected to one of the clinical pathways and
#the other to both of them.

data(GenPhen)
W0 <- matrix(c(rep(1,2),rep(0,8),rep(1,2),rep(0,8),rep(1,3),rep(0,7),rep(1,2)),nrow=8,ncol=4)
B0 <- matrix(c(rep(0,8),rep(1,2),rep(0,3),1,rep(0,2)),nrow=4,ncol=4)

#Estimation only
GSCA(GenPhen,W0, B0,estim=TRUE,path.test=FALSE)
#Estimation and test for all the path coefficients in the model
GSCA(GenPhen,W0, B0,estim=TRUE,path.test=TRUE)
#Test only
GSCA(GenPhen,W0, B0,estim=FALSE,path.test=TRUE)
#Give names to the latent variables
GSCA(GenPhen,W0, B0,latent.names=c("Gene1","Gene2","Clinical pathway 1","Clinical pathway 2"),
estim=TRUE,path.test=TRUE)
#Testing only a subset of path coefficients
GSCA(GenPhen,W0, B0,estim=FALSE,path.test=TRUE,path=matrix(c(1,2,3,4),ncol=2))
```

`GSCAestim` *Structural Equation Models for multiple genotypes and multiple traits using Generalized Structured Component Analysis.*

Description

GSCAestim fits the Generalized Structured Component Analysis (GSCA) model to data on multiple genetic variants and multiple traits (see Romdhani et al., 2014). An Alternating Least-Squares algorithm (ALS) (de Leeuw, Young and Takane, 1976) is used to minimize a global least squares criterion. The ALS algorithm alternates between two main steps until convergence. In the first step, the weight coefficients are fixed, and the path coefficients are updated in the least-squares sense. In the second step, the weights are updated in the least-squares sense for fixed path coefficients.

Usage

```
GSCAestim(data,W0,B0)
```

Arguments

data	Data frame containing the observed variables.
W0	Matrix with 0's and 1's indicating connections between the observed variables (genotypes and traits) and the latent variables (genes and clinical pathways). The rows correspond to the observed variables in the same order as in data; the columns to the latent variables. A value of 1 indicates an arrow from the observed variable in the row to the latent variable in the column.
B0	square matrix with 0's and 1's indicating connections among latent variables (genes and clinical pathways). Both rows and columns correspond to the latent variables. A value of 1 indicates an arrow directed from the latent variable in the row to the latent variable in the column.

Value

Returns a list with 2 items.

Weight	Matrix of the same dimension as W0 with 1's replaced by weight coefficients estimates.
Path	Matrix of the same dimension as B0 with 1's replaced by path coefficients estimates.

Author(s)

Hela Romdhani, Stepan Grinek, Heungsun Hwang and Aurelie Labbe

References

de Leeuw, J., Young, F. W., and Takane, Y. (1976). Additive structure in qualitative data: An alternating least squares method with optimal scaling features. *Psychometrika*, 41, 471-503.

Romdhani, H., Hwang, H., Paradis, G., Roy-Gagnon, M.-H. and Labbe, A. (2014). Pathway-based Association Study of Multiple Candidate Genes and Multiple Traits Using Structural Equation Models. Submitted.

Examples

```
#Scenario (g) in Romdhani et al. (2014): 4 SNPs mapped to 2 genes
#and 4 traits involved in 2 clinical pathways.
#In total: 8 observed variables and 4 latent variables.
#One of the traits is involved in both clinical pathways.
#One gene is connected to one of the clinical pathways and the other
#to both of them.
data(GenPhen)
W0 <- matrix(c(rep(1,2),rep(0,8),rep(1,2),rep(0,8),rep(1,3),rep(0,7),rep(1,2)),nrow=8,ncol=4)
B0 <- matrix(c(rep(0,8),rep(1,2),rep(0,3),1,rep(0,2)),nrow=4,ncol=4)
res<-GSCAestim(data=GenPhen,W0,B0)
```

QCAHS *Dataset used in the vignette.*

Description

Dataset contaning some variables of interest from the Quebec Child and Adolescent Health and Social Survey (QCAHS), observed on 1707 French Canadian participants (860 boys and 847 girls). Detailed descriptions of the QCAHS design and methods can be found in Paradis et al. (2003). The dataset contains 8 traits (z-score transformation standardized for age and sex), 33 SNPs and 2 polymorphisms with more than two alleles.

Usage

```
data(GenPhen)
```

Format

A data frame of 49 columns and 1707 rows.

References

Romdhani, H., Hwang, H., Paradis, G., Roy-Gagnon, M.-H. and Labbe, A. (2014). Pathway-based Association Study of Multiple Candidate Genes and Multiple Traits Using Structural Equation Models, submitted. Paradis, G., Lambert, M., O'Loughlin, J., Lavallee, C., Aubin, J., Berthiaume, P., Ledoux, M., Delvin, E., Levy, E., and Hanley, J. (2003). The quebec child and adolescent health and social survey: design and methods of a cardiovascular risk factor survey for youth. Can J Cardiol, 19:523-531.

ResQCAHS *A list containing the results obtained for the QCAHS dataset.*

Description

A list of 3 matrices: Weight containing the weight estimates, Path containing path coefficients estimates and pvalues containing p-values for all path coefficients.

Usage

```
data(ResQCAHS)
```

Format

A list of 3 matrices.

W0 *Matrix indicating connections between the observed variables and the latent variables for QCAHS data.*

Description

Matrix indicating connections between the observed variables and the latent variables in the path model fitted to QCAHS data, used in the vignette to estimate the weight coefficients of the model.

Usage

```
data(W0)
```

Format

A matrix of 49 columns and 28 rows.

Index

∗ **Path matrix**
B0, 3
∗ **Real data results.**
ResQCAHS, 7
∗ **Weight matrix**
W0, 8
∗ **datasets**
GenPhen, 3
QCAHS, 7

ASGSCA *(*ASGSCA-package*)*, 2
ASGSCA-package, 2

B0, 3

GenPhen, 3
GSCA, 4
GSCAestim, 5

QCAHS, 7

ResQCAHS, 7

W0, 8