Package ‘o2plsda’

July 19, 2024

Type Package

Title Multiomics Data Integration

Version 0.0.25

Description Provides functions to do ‘O2PLS-DA’ analysis for multiple omics data integration. The algorithm came from ‘O2-PLS, a two-block (X±Y) latent variable regression (LVR) method with an integral OSC filter’ which published by Johan Trygg and Svante Wold at 2003 <doi:10.1002/cem.775>. 'O2PLS' is a bidirectional multivariate regression method that aims to separate the covariance between two data sets (it was recently extended to multiple data sets) (Löfstedt and Trygg, 2011 <doi:10.1002/cem.1388>; Löfstedt et al., 2012 <doi:10.1016/j.aca.2013.06.026>) from the systematic sources of variance being specific for each data set separately.

License GPL-3

Imports Rcpp (>= 1.0.7), dplyr, magrittr, parallel, ggplot2, ggrepel, methods, stats

Encoding UTF-8

Suggests knitr, markdown, rmarkdown

VignetteBuilder knitr

LinkingTo Rcpp, RcppArmadillo

RoxygenNote 7.3.1

NeedsCompilation yes

Repository CRAN

Author Kai Guo [aut, cre],
Junguk Hur [aut],
Eva Feldman [aut]

Maintainer Kai Guo <guokai8@gmail.com>

Date/Publication 2024-07-18 22:40:02 UTC
loadings ............................ 2
loadings.o2plsdab ........................ 3
loadings.pldsab ........................ 3
o2cv .............................. 4
o2pls ................................ 5
O2pls-class .......................... 7
oplsdal ................................ 7
plot.O2plsa .......................... 8
plot.o2plsdab ........................ 10
plot.plsdada .......................... 11
plsdada ............................. 12
print.O2plsa .......................... 14
print.plsdada .......................... 14
scores ................................ 15
scores.O2plsd ....... ........................ 15
scores.o2plsdada .......................... 16
scores.plsdada .................................. 16
summary.O2plsd .................................. 17
summary.plsdada .................................. 18
vip .................................. 18

Index .......................... 20

loadings ............................ 2
Extract the loadings from an O2PLS fit

Description
This function extracts loading parameters from an O2PLS fit
This function extracts loading parameters from an O2PLS fit

Usage
loadings(x, ...)

## S3 method for class 'O2plsd'
loadings(x, loading = c("Xjoint", "Yjoint", "Xorth", "Yorth"), ...)

Arguments

x  Object of class O2plsd
...
loading  the loadings for one of "Xjoint", "Yjoint", "Xorth", "Yorth"
loadings.o2plsda

Value
Loading matrix
Loading matrix

loadings.o2plsda extract the loading value from the O2PLSDA analysis

Description
extract the loading value from the O2PLSDA analysis

Usage
## S3 method for class 'o2plsda'
loadings(x, loading = "Xloading", ...)

Arguments
x Object of class o2plsda
loading the loadings for one of "Xjoint", "Yjoint", "Xorth", "Yorth"
... For consistency

loadings.plsda extract the loading value from the PLSDA analysis

Description
extract the loading value from the PLSDA analysis

Usage
## S3 method for class 'plsda'
loadings(x, ...)

Arguments
x Object of class plsda
... For consistency
Cross validation for O2PLS

Usage

```r
o2cv(
  X,
  Y,
  nc,
  nx,
  ny,
  group = NULL,
  nr_folds = 5,
  ncores = 1,
  scale = FALSE,
  center = FALSE
)
```

Arguments

- `X`: a Numeric matrix (input)
- `Y`: a Numeric matrix (input)
- `nc`: Integer. Number of joint PLS components.
- `nx`: Integer. Number of orthogonal components in X
- `ny`: Integer. Number of orthogonal components in Y
- `group`: a vector to indicate the group for Y
- `nr_folds`: Integer to indicate the folds for cross validation
- `ncores`: Integer. Number of CPUs to use for cross validation
- `scale`: boolean values determining if data should be scaled or not
- `center`: boolean values determining if data should be centered or not

Value

- a data frame with the Q and RMSE values

Author(s)

Kai Guo
Examples

```r
set.seed(123)
X = matrix(rnorm(500),50,10)
Y = matrix(rnorm(500),50,10)
X = scale(X, scale = TRUE)
Y = scale(Y, scale = TRUE)
# group factor could be omitted if you don't have any group
group <- rep(c("Ctrl","Treat"), each = 25)
cv <- o2cv(X, Y, 1:2, 1:2, 1:2, group=group, nr_folds = 2, ncores=1)
```

Description

fit O2PLS model with best nc, nx, ny

Usage

```r
o2pls(X, Y, nc, nx, ny, scale = FALSE, center = FALSE)
```

Arguments

- `X`: a Numeric matrix (input)
- `Y`: a Numeric matrix (input)
- `nc`: Integer. Number of joint PLS components.
- `nx`: Integer. Number of orthogonal components in X
- `ny`: Integer. Number of orthogonal components in Y
- `scale`: boolean values determining if data should be scaled or not
- `center`: boolean values determining if data should be centered or not

Value

An object containing

- `Xscore`: Joint X scores
- `Xloading`: Joint X loadings
- `Yscore`: Joint Y scores
- `Yloading`: Joint Y loadings
- `TYosc`: Orthogonal X scores
- `PYosc`: Orthogonal X loadings
- `WYosc`: Orthogonal X weights
- `UXosc`: Orthogonal Y scores
- `PXosc`: Orthogonal Y loadings
CXosc Orthogonal Y weights
BU Regression coefficient in $T_t \sim U$
BT Regression coefficient in $U \sim T_t$
Xhat Prediction of $X$ with $Y$
Yhat Prediction of $Y$ with $X$
R2Xhat Variation of the predicted $X$ as proportion of variation in $X$
R2Yhat Variation of the predicted $Y$ as proportion of variation in $Y$
R2X Variation of the modeled part in $X$ (defined by Joint + Orthogonal variation) as proportion of total variation in $X$
R2Y Variation of the modeled part in $Y$ (defined by Joint + Orthogonal variation) as proportion of total variation in $Y$
R2Xcorr Variation of the joint part in $X$
R2Ycorr Variation of the joint part in $Y$
R2Xo Variation of the orthogonal part in $X$ as proportion of variation in $X$
R2Yo Variation of the orthogonal part in $Y$ as proportion of variation in $Y$
R2Xp Variation in $X$ joint part predicted by $Y$ Joint part
R2Yp Variation in $Y$ joint part predicted by $X$ Joint part
varXj Variation in each Latent Variable (LV) in $X$ Joint part
varYj Variation in each Latent Variable (LV) in $Y$ Joint part
varXorth Variation in each Latent Variable (LV) in $X$ Orthogonal part
varYorth Variation in each Latent Variable (LV) in $Y$ Orthogonal part
Exy Residuals in $X$
Fxy Residuals in $Y$

Author(s)
Kai Guo

Examples

```r
set.seed(123)
X = matrix(rnorm(500),50,10)
Y = matrix(rnorm(500),50,10)
X = scale(X, scale = TRUE)
Y = scale(Y, scale = TRUE)
fit <- o2pls(X, Y, 1, 2, 2)
summary(fit)
```
Class "O2pls" This class represents the Annotation information

Slots

- **X**: A Numeric matrix (input)
- **Y**: A Numeric matrix (input)
- **params**: Parameters used in O2pls analysis
- **results**: List of O2pls results

Author(s)

Kai Guo

---

Class "O2pls" This class represents the Annotation information

**Description**

Computes orthogonal scores partial least squares regressions with the NIPALS algorithm. It returns a comprehensive set of pls outputs (e.g. scores and vip).

**Usage**

```r
oplsda(X, Y, nc, scale = FALSE, center = TRUE, maxiter = 100, tol = 1e-05)
```

**Arguments**

- **X**: A O2pls object or a matrix of predictor variables.
- **Y**: A single vector indicating the group
- **nc**: The number of pls components (the one joint components + number of orthogonal components).
- **scale**: Logical indicating whether X must be scaled (suggest TRUE).
- **center**: Boolean values determining if data should be centered or not.
- **maxiter**: Maximum number of iterations.
- **tol**: Limit for convergence of the algorithm in the Nipals algorithm.
Value

a list containing the following elements:

- nc the number of components used (one joint component + number of orthogonal components
- scores a matrix of scores corresponding to the observations in X. The components retrieved correspond to the ones optimized or specified.
- Xloadings a matrix of loadings corresponding to the explanatory variables. The components retrieved correspond to the ones optimized or specified.
- Yloadings a matrix of partial least squares loadings corresponding to Y
- vip the VIP matrix.
- xvar a matrix indicating the standard deviation of each component (sd), the variance explained by each single component (explained_var) and the cumulative explained variance (cumulative_explained_var). These values are computed based on the data used to create the projection matrices.
- projection_matrix the matrix of projection matrix
- weight a matrix of partial least squares ("pls") weights.

Author(s)

Kai Guo

Examples

X <- matrix(rnorm(50),10,5)
Y <- matrix(rnorm(50),10,5)
fit <- o2pls(X,Y,2,1,1)
yy <- rep(c(0,1),5)
fit0 <- oplsda(fit,yy,2)

plot.O2pls  
Score or loading plot for the O2PLS results

Description

Score or loading plot for the O2PLS results

Usage

## S3 method for class 'O2pls'
plot(
x,
type = "score",
var = "Xjoint",
group = NULL,
ind = c(1, 2),

plot.02pls
plot.O2pls

```r
color = NULL,
top = 20,
ellipse = TRUE,
order = FALSE,
pt.size = 3,
label = TRUE,
label.size = 4,
repel = TRUE,
rotation = FALSE,
...)
```

Arguments

- **x**: an O2pls object
- **type**: score or loading
- **var**: specify Xjoint
- **group**: color used for score plot
- **ind**: which components to be used for score plot or loading plot
- **color**: color used for score or loading plot
- **top**: the number of largest loading value to plot
- **ellipse**: TRUE/FALSE
- **order**: order by the value or not
- **pt.size**: point size
- **label**: plot label or not (TRUE/FALSE)
- **label.size**: label size
- **repel**: use ggrepel to show the label or not
- **rotation**: flip the figure or not (TRUE/FALSE)
- ... For consistency

Value

a ggplot2 object

Author(s)

Kai Guo

Examples

```r
X <- matrix(rnorm(50),10,5)
Y <- matrix(rnorm(50),10,5)
fit <- o2pls(X,Y,2,1,1)
plot(fit, type="score")
```
plot.o2plsd

Score, VIP or loading plot for the O2PLS results

Description
Score, VIP or loading plot for the O2PLS results

Usage
## S3 method for class 'o2plsd'
plot(
  x,
  type = "score",
  group = NULL,
  ind = c(1, 2),
  color = NULL,
  top = 20,
  ellipse = TRUE,
  order = FALSE,
  pt.size = 3,
  label = TRUE,
  label.size = 4,
  repel = FALSE,
  rotation = FALSE,
  ...
)

Arguments

x an o2plsd object
type score, vip or loading
group color used for score plot
ind which components to be used for score plot or loading plot
color color used for score or loading plot
top the number of largest loading value to plot
ellipse TRUE/FALSE
order order by the value or not
pt.size point size
label plot label or not (TRUE/FALSE)
label.size label size
repel use ggrepel to show the label or not
rotation flip the figure or not (TRUE/FALSE)
... For consistency
Value

a ggplot2 object

Author(s)

Kai Guo

Examples

```r
X <- matrix(rnorm(50),10,5)
Y <- matrix(rnorm(50),10,5)
fit <- o2pls(X,Y,2,1,1)
yy <- rep(c(0,1),5)
fit0 <- oplsda(fit,yy,2)
plot(fit0, type="score", group = factor(yy))
```

Description

Score, VIP or loading plot for the plsda results

Usage

```r
## S3 method for class 'plsda'
plot(
  x,
  type = "score",
  group = NULL,
  ind = c(1, 2),
  color = NULL,
  top = 20,
  ellipse = TRUE,
  order = FALSE,
  pt.size = 3,
  label = TRUE,
  label.size = 4,
  repel = FALSE,
  rotation = FALSE,
  ...
)
```
Arguments

x an plsda object

type score, vip or loading

group color used for score plot

ind which components to be used for score plot or loading plot

color color used for score or loading plot

top the number of largest loading value to plot

ellipse TRUE/FALSE

order order by the value or not

pt.size point size

label plot label or not (TRUE/FALSE)

label.size label size

repel use ggrepel to show the label or not

rotation flip the figure or not (TRUE/FALSE)

... For consistency

Value

a ggplot2 object

Author(s)

Kai Guo

Examples

X <- matrix(rnorm(500),10,50)
Y <- rep(c("a","b"),each=5)
fit0 <- plsda(X,Y,2)
plot(fit0, type = "score", group = factor(Y))

pllda

Partial least squares discriminant analysis

Description

Perform a PLS discriminant analysis

Usage

pllda(X, Y, nc, scale = TRUE, center = TRUE, cv = TRUE, nr_folds = 5)
Arguments

\(X\) a matrix of predictor variables.
\(Y\) a single vector indicating the group
\(nc\) the number of pls components (the one joint components + number of orthogonal components).
\(scale\) logical indicating whether \(X\) must be scaled (suggest TRUE).
\(center\) logical indicating whether \(X\) must be centered (suggest TRUE).
\(cv\) logical indicating whether cross-validation will be performed or not (suggest TRUE).
\(nr\_folds\) \(nr\_folds\) Integer to indicate the folds for cross validation.

Value

a list containing the following elements:

- \(nc\) the number of components used (one joint components + number of orthogonal components)
- \(scores\) a matrix of scores corresponding to the observations in \(X\), The components retrieved correspond to the ones optimized or specified.
- \(Xloadings\) a matrix of loadings corresponding to the explanatory variables. The components retrieved correspond to the ones optimized or specified.
- \(vip\) the VIP matrix.
- \(xvar\) variance explained of \(X\) by each single component.
- \(R2Y\) variance explained of \(Y\) by each single component.
- \(PRESS\) The residual sum of squares for the samples which were not used to fit the model
- \(Q2\) quality of cross-validation

Author(s)

Kai Guo

Examples

\(X \leftarrow \text{matrix(rnorm(500),10,50)}\)
\(Y \leftarrow \text{rep(c("a","b"),each=5)}\)
\(fit \leftarrow \text{plsda(X,Y,2)}\)
**print.02pls**  
*Print the summary of O2PLS results.*

**Description**  
Print the summary of O2PLS results.

**Usage**  
```r
## S3 method for class '02pls'
print(x, ...)
```

**Arguments**  
- `x`: An O2pls object  
- `...`: For consistency

**Author(s)**  
Kai Guo

**Examples**  
```r
X <- matrix(rnorm(50),10,5)
Y <- matrix(rnorm(50),10,5)
object <- o2pls(X,Y,1,1,1)
print(object)
```

---

**print.plsda**  
*Print the summary of plsda results.*

**Description**  
Print the summary of plsda results.

**Usage**  
```r
## S3 method for class 'plsda'
print(x, ...)
```

**Arguments**  
- `x`: An plsda object  
- `...`: For consistency

**Examples**  
```r
X <- matrix(rnorm(50),10,5)
Y <- matrix(rnorm(50),10,5)
object <- plsda(X,Y,1,1,1)
print(object)
```
scores

Author(s)

Kai Guo

Examples

X <- matrix(rnorm(500),10,50)
Y <- rep(c("a","b"),each=5)
fit <- plsda(X,Y,2)
print(fit)

---------

scores Extract the scores from an O2PLS fit

Description

This function extracts score matrices from an O2PLS fit

Usage

scores(x, ...)

Arguments

x Object of class O2pls
...
For consistency

Value

Scores matrix

---------

scores.O2pls Extract the scores from an O2PLS fit

Description

This function extracts scores parameters from an O2PLS fit

Usage

## S3 method for class 'O2pls'
scores(x, score = c("Xjoint", "Yjoint", "Xorth", "Yorth"), ...)

Arguments

x Object of class O2pls
score the scores matrix for one of "Xjoint", "Yjoint", "Xorth", "Yorth"
...
Other arguments
scores.o2plsda

Extract the scores from an O2PLS DA analysis

Description
Extract the scores from an O2PLS DA analysis

Usage
## S3 method for class 'o2plsda'
scores(x, ...)

Arguments
x Object of class o2plsda
...

Value
score matrix

Author(s)
Kai Guo

scores.plsda

Extract the scores PLSDA analysis

Description
Extract the scores PLSDA analysis

Usage
## S3 method for class 'plsda'
scores(x, ...)

Arguments
x Object of class plsda
...

Value
score matrix

Author(s)
Kai Guo
Summary of an O2PLS object

Description
Summary of an O2PLS object

Usage
## S3 method for class 'O2pls'
summary(object, ...)

Arguments

object a O2pls object
...

For consistency

Value
Detail of O2PLS results

Author(s)
Kai Guo

Examples
X <- matrix(rnorm(50),10,5)
Y <- matrix(rnorm(50),10,5)
object <- o2pls(X,Y,1,1,1)
summary(object)
### summary.plsda

**Summary of an plsda object**

#### Description

Summary of an plsda object

#### Usage

```r
## S3 method for class 'plsda'
summary(object, ...)
```

#### Arguments

- `object`: a plsda object
- `...`: For consistency

#### Value

Detail of plsda results

#### Author(s)

Kai Guo

#### Examples

```r
X <- matrix(rnorm(500),10,50)
Y <- rep(c("a","b"),each=5)
fit <- plsda(X,Y,2)
summary(fit)
```

### vip

**Extract the VIP values from the O2PLS-DA object**

#### Description

Extract the VIP values from the O2PLS-DA object

#### Usage

```r
tip(x)
```

#### Arguments

- `x`: the o2plsa object or plsda object
<table>
<thead>
<tr>
<th>vip</th>
<th>19</th>
</tr>
</thead>
</table>

**Value**
- a data frame
Index

* classes
  O2pls-class, 7

loadings, 2
loadings.o2plsda, 3
loadings.plsda, 3

o2cv, 4
o2pls, 5
O2pls-class, 7
oplsda, 7

plot.O2pls, 8
plot.o2plsda, 10
plot.plsda, 11
plsda, 12
print.O2pls, 14
print.plsda, 14

scores, 15
scores.O2pls, 15
scores.o2plsda, 16
scores.plsda, 16
summary.O2pls, 17
summary.plsda, 18

vip, 18