Package: guidedPLS (via r-universe)

September 12, 2024

Type Package

Title Supervised Dimensional Reduction by Guided Partial Least Squares

Version 0.99.0

Depends R (>= 3.4.0)

Imports irlba

Suggests fields, knitr, rmarkdown, testthat

Description Guided partial least squares (guided-PLS) is the combination of partial least squares by singular value decomposition (PLS-SVD) and guided principal component analysis (guided-PCA). For the details of the methods, see the reference section of GitHub README.md <https://github.com/rikenbit/guidedPLS>.

License MIT + file LICENSE

URL `https://github.com/rikenbit/guidedPLS`

VignetteBuilder knitr

Repository https://rikenbit.r-universe.dev

RemoteUrl https://github.com/rikenbit/guidedpls

RemoteRef HEAD

RemoteSha a0875b7772a54fa7d8eed2df2d98691f6344d098

Contents

guidedPLS-package *Supervised Dimensional Reduction by Guided Partial Least Squares*

Description

Guided partial least squares (guided-PLS) is the combination of partial least squares by singular value decomposition (PLS-SVD) and guided principal component analysis (guided-PCA). For the details of the methods, see the reference section of GitHub README.md <https://github.com/rikenbit/guidedPLS>.

Details

The DESCRIPTION file: This package was not yet installed at build time.

Index: This package was not yet installed at build time.

Author(s)

Koki Tsuyuzaki [aut, cre]

Maintainer: Koki Tsuyuzaki <k.t.the-answer@hotmail.co.jp>

References

Le Cao, et al. (2008). A Sparse PLS for Variable Selection when Integrating Omics Data. *Statistical Applications in Genetics and Molecular Biology*, 7(1)

Reese S E, et al. (2013). A new statistic for identifying batch effects in high-throughput genomic data that uses guided principal component analysis. *Bioinformatics*, 29(22), 2877-2883

See Also

`toyModel`,`PLSSVD`,`sPLSDA`,`guidedPLS`

Examples

```
ls("package:guidedPLS")
```

`dummyMatrix` *Toy model data for using dNMF, dSVD, dsiNMF, djNMF, dPLS, dNTF, and dNTD*

Description

A label vector is converted to a dummy matrix.

Usage

```
dummyMatrix(y, center=TRUE)
```

Arguments

`y` A label vector to specify the group of data.

`center` An option to center the rows of matrix (Default: TRUE).

Value

A matrix is generated. The number of row is equal to the length of y and the number of columns is the number of unique elements of y.

Author(s)

Koki Tsuyuzaki

Examples

```
y <- c(1, 3, 2, 1, 4, 2)
dummyMatrix(y)
```

`guidedPLS` *Guided Partial Least Squares (guied-PLS)*

Description

Four matrices X1, X2, Y1, and Y2 are required. X1 and Y1 are supposed to share the rows, X2 and Y2 are supposed to share the rows, and Y1 and Y2 are supposed to share the columns.

Usage

```
guidedPLS(X1, X2, Y1, Y2, k=.minDim(X1, X2, Y1, Y2),
    cortest=FALSE, fullrank=TRUE, verbose=FALSE)
```

Arguments

X1 The input matrix which has N-rows and M-columns.

Y1 The input matrix which has N-rows and L-columns.

X2 The input matrix which has O-rows and P-columns.

Y2 The input matrix which has O-rows and L-columns.

k The number of low-dimension (k < N, M, L, O, Default: .minDim(X1, X2, Y1, Y2))

cortest If cortest is set as TRUE, t-test of correlation coefficient is performed (Default: FALSE)

fullrank If fullrank is set as TRUE, irlba is used, otherwise fullrank SVD is used (Default: TRUE)

verbose Verbose option (Default: FALSE)

Value

res: object of svd() loadingYX1: Loading vector to project X1 to lower dimension via Y1 (M times k). loadingYX2: Loading vector to project X2 to lower dimension via Y2 (P times k). scoreX1: Projected X1 (N times k) scoreX2: Projected X2 (O times k) scoreYX1: Projected YX1 (L times k) scoreYX2: Projected YX2 (L times k) corYX1: Correlation Coefficient (Default: NULL) corYX2: Correlation Coefficient (Default: NULL) pvalYX1: P-value vector of corYX1 (Default: NULL) pvalYX2: P-value vector of corYX2 (Default: NULL) qvalYX1: Q-value vector of BH method against pvalYX1 (Default: NULL) qvalYX2: Q-value vector of BH method against pvalYX2 (Default: NULL)

Author(s)

Koki Tsuyuzaki

References

Le Cao, et al. (2008). A Sparse PLS for Variable Selection when Integrating Omics Data. *Statistical Applications in Genetics and Molecular Biology*, 7(1)

Reese S E, et al. (2013). A new statistic for identifying batch effects in high-throughput genomic data that uses guided principal component analysis. *Bioinformatics*, 29(22), 2877-2883

Examples

```
# Test data
data <- toyModel()

# Simple usage
out <- guidedPLS(X1=data$X1, X2=data$X2, Y1=data$Y1, Y2=data$Y2, k=4)
```

PLSSVD *Partial Least Squares by Singular Value Decomposition (PLS-SVD)*

Description

Two matrices X and Y sharing a row are required

Usage

```
PLSSVD(X, Y, k=.minDim(X, Y), deflation=FALSE, fullrank=TRUE, verbose=FALSE)
```

Arguments

`X`	The input matrix which has N-rows and M-columns.
`Y`	The input matrix which has N-rows and L-columns.
`k`	The number of low-dimension (k < N, M, L, Default: .minDim(X, Y))
`deflation`	If deflation is set as TRUE, the score vectors are made orthogonal, otherwise the loading vectors are made orthogonal (Default: FALSE)
`fullrank`	If fullrank is set as TRUE, irlba is used, otherwise fullrank SVD is used (Default: TRUE)
`verbose`	Verbose option (Default: FALSE)

Value

scoreX : Score matrix which has M-rows and K-columns. loadingX : Loading matrix which has N-rows and K-columns. scoreY : Score matrix which has L-rows and K-columns. loadingY : Loading matrix which has N-rows and K-columns. d : K-length singular value vector of the cross-product matrix X'Y.

Author(s)

Koki Tsuyuzaki

References

Le Cao, et al. (2008). A Sparse PLS for Variable Selection when Integrating Omics Data. *Statistical Applications in Genetics and Molecular Biology*, 7(1)

Examples

```
# Test data
data <- toyModel()

# Simple usage
out <- PLSSVD(X=data$X1, Y=data$Y1, k=4)
```

softThr *Soft-thresholding to make a sparse vector sparse*

Description

The degree of the sparseness of vector is controlled by the lambda parameter.

Usage

```
softThr(y, lambda=1)
```

Arguments

`y` A numerical vector.

`lambda` Threshold value to convert a value 0. If the absolute value of an element of vector is less than lambda, the value is converted to 0 (Default: 1).

Value

A numerical vector, whose length is the same as that of y.

Author(s)

Koki Tsuyuzaki

Examples

```
y <- seq(-2, 2, 0.1)
softThr(y)
```

sPLSDA *Sparse Partial Least Squares Discriminant Analysis (sPLS-DA)*

Description

Two matrices X and Y sharing a row are required

Usage

```
sPLSDA(X, Y, k=.minDim(X, Y), lambda=1, thr=1e-10, fullrank=TRUE,
    num.iter=10, verbose=FALSE)
```

Arguments

`X` The input matrix which has N-rows and M-columns.

`Y` The input matrix which has N-rows and L-columns.

`k` The number of low-dimension (k < N, M, L, Default: .minDim(X, Y))

`lambda` Penalty parameter to control the sparseness of u and v. The larger the value, the sparser the solution (Default: 1).

`thr` Threshold to stop the iteration (Default: 1e-10).

`fullrank` If fullrank is set as TRUE, irlba is used, otherwise fullrank SVD is used (Default: TRUE)

`num.iter` The number of iterations in each rank (Default: 10)

`verbose` Verbose option (Default: FALSE)

Value

scoreX : Score matrix which has M-rows and K-columns. loadingX : Loading matrix which has N-rows and K-columns. scoreY : Score matrix which has L-rows and K-columns. loadingY : Loading matrix which has N-rows and K-columns. d : K-length singular value vector of the cross-product matrix X'Y.

Author(s)

Koki Tsuyuzaki

References

Le Cao, et al. (2008). A Sparse PLS for Variable Selection when Integrating Omics Data. *Statistical Applications in Genetics and Molecular Biology*, 7(1)

Examples

```
# Test data
data <- toyModel()

# Simple usage
out <- sPLSDA(X=data$X1, Y=data$Y1, k=4)
```

`toyModel` *Toy model data for using PLSSVD, sPLSDA, and guidedPLS*

Description

The data is used for confirming the algorithm are properly working.

Usage

```
toyModel(model=”Easy”, seeds=123)
```

Arguments

`model`	"Easy" and "Hard" are available (Default: "Easy").
`seeds`	Random number for setting set.seeds in the function (Default: 123).

Value

A list object containing a set of matrices X1, X2, Y1, Y1_dummy, Y2, Y1_dummy.

Author(s)

Koki Tsuyuzaki

See Also

`PLSSVD`,`sPLSDA`,`guidedPLS`

Examples

```
data <- toyModel(seeds=123)
```

Index