Package 'msPCA'

December 9, 2025

Type Package

Title Sparse Principal Component Analysis with Multiple Principal Components

Version 0.1.0

Date 2025-12-02

Description Implements an algorithm for computing multiple sparse principal components of a dataset. The method is based on Cory-Wright and Pauphilet ``Sparse PCA with Multiple Principal Components" (2022) <doi:10.48550/arXiv.2209.14790>. The algorithm uses an iterative deflation heuristic with a truncated power method applied at each iteration to compute sparse principal components with controlled sparsity.

License MIT + file LICENSE

Imports Rcpp (>= 1.0.11)

LinkingTo Rcpp, RcppEigen

RoxygenNote 7.3.3

Encoding UTF-8

NeedsCompilation yes

Author Ryan Cory-Wright [aut, cph] (ORCID: <https://orcid.org/0000-0002-4485-0619>), Jean Pauphilet [aut, cre, cph] (ORCID: <https://orcid.org/0000-0001-6352-0984>)

Maintainer Jean Pauphilet <jpauphilet@london.edu>

Repository CRAN

Date/Publication 2025-12-09 08:50:02 UTC

Contents

fraction_variance_explained 2
fraction_variance_explained_perPC 2
mspca 3
orthogonality_violation 4

print_mspca 5
tpw 5
variance_explained_perPC 6

Index **7**

`fraction_variance_explained` *Fraction of variance explained*

Description

Computes the fraction of variance explained (variance explained normalized by the trace of the covariance/correlation matrix) by a set of PCs.

Usage

```
fraction_variance_explained(C, U)
```

Arguments

`C`	A matrix. The correlation or covariance matrix (p x p).
`U`	A matrix. The matrix containing the r PCs (p x r).

Value

A float.

Examples

```
library(datasets)
TestMat <- cor(datasets::mtcars)
mspcares <- mspca(TestMat, 2, c(4,4))
fraction_variance_explained(TestMat, mspcares$x_best)
```

`fraction_variance_explained_perPC` *Fraction of variance explained per PC*

Description

Computes the fraction of variance explained (variance explained normalized by the trace of the covariance/correlation matrix) by each PC.

Usage

```
fraction_variance_explained_perPC(C, U)
```

Arguments

C	A matrix. The correlation or covariance matrix (p x p).
U	A matrix. The matrix containing the r PCs (p x r).

Value

An array.

---

mspca *Multiple Sparse PCA*

Description

Returns multiple sparse principal component of a matrix using an iterative deflation heuristic.

Usage

```
mspca(
  Sigma,
  r,
  ks,
  maxIter = 200L,
  verbose = TRUE,
  violationTolerance = 1e-04,
  stallingTolerance = 1e-08,
  maxIterTPW = 200L,
  timeLimitTPW = 20L
)
```

Arguments

Sigma	A matrix. The correlation or covariance matrix, whose sparse PCs will be computed.
r	An integer. Number of principal components (PCs) to be computed.
ks	A list of integers. Target sparsity of each PC.
maxIter	(optional) An integer. Maximum number of iterations of the algorithm. Default 200.
verbose	(optional) A Boolean. Controls console output. Default TRUE.
violationTolerance	(optional) A float. Tolerance for the violation of the orthogonality constraints. Default 1e-4
stallingTolerance	(optional) A float. Controls the objective improvement below which the algorithm is considered to have stalled. Default 1e-8

`maxIterTPW`	(optional) An integer. Maximum number of iterations of the truncated power method (inner iteration). Default 200.
`timeLimitTPW`	(optional) An integer. Maximum time in seconds for the truncated power method (inner iteration). Default 20.

Value

An object with 4 fields: ‘x_best‘ (p x r array containing the sparse PCs), ‘objective_value‘, ‘orthogonality_violation‘, ‘runtime‘.

Examples

```
library(datasets)
TestMat <- cor(datasets::mtcars)
mspca(TestMat, 2, c(4,4))
```

`orthogonality_violation` *Orthogonality constraint violation*

Description

Computes the orthogonality constraint violation defined as the distance (infinity norm) between $U^\top U$ and the identity matrix.

Usage

```
orthogonality_violation(U)
```

Arguments

`U`	A matrix. Each column correspond to an p-dimensional PC.

Value

A float.

Examples

```
library(datasets)
TestMat <- cor(datasets::mtcars)
mspcares <- mspca(TestMat, 2, c(4,4))
orthogonality_violation(mspcares$x_best)
```

print_mspca *Print mspca output*

Description

Displays the output of the msPCA algorithm.

Usage

```
print_mspca(sol_object, C)
```

Arguments

sol_object	A list. The output of the mspca or twp function.
C	A matrix. The correlation or covariance matrix (p x p).

Value

None. Prints output to console.

Examples

```
library(datasets)
TestMat <- cor(datasets::mtcars)
mspcares <- mspca(TestMat, 2, c(4,4))
print_mspca(mspcares, TestMat)
```

tpw *Truncated Power Method*

Description

Returns the leading sparse principal component of a matrix using the truncated power method.

Usage

```
tpw(Sigma, k, maxIter = 200L, verbose = TRUE, timeLimit = 10L)
```

Arguments

Sigma	A matrix. The correlation or covariance matrix, whose sparse PCs will be computed.
k	An integer. Target sparsity of the PC.
maxIter	(optional) An integer. Maximum number of iterations of the algorithm. Default 200.
verbose	(optional) A Boolean. Controls console output. Default TRUE.
timeLimit	(optional) An integer. Maximum time in seconds. Default 10.

Value

An object with 3 fields: ‘x_best‘ (p x 1 array containing the sparse PC), ‘objective_value‘, ‘runtime‘.

References

Yuan, X. T., & Zhang, T. (2013). Truncated power method for sparse eigenvalue problems. The Journal of Machine Learning Research, 14(1), 899-925.

Examples

```
library(datasets)
TestMat <- cor(datasets::mtcars)
tpw(TestMat, 4)
```

`variance_explained_perPC` *Variance explained per PC*

Description

Computes the variance explained by each PC.

Usage

```
variance_explained_perPC(C, U)
```

Arguments

`C`	A matrix. The correlation or covariance matrix (p x p).
`U`	A matrix. The matrix containing the r PCs (p x r).

Value

An array.

Index

fraction_variance_explained, 2
fraction_variance_explained_perPC, 2

mspca, 3

orthogonality_violation, 4

print_mspca, 5

tpw, 5

variance_explained_perPC, 6