Package ‘PCA4you’

October 22, 2017

Type Package

Title PCA Plotting in a User Friendly Way

Version 1.2.4

Maintainer Martin Berntsson <`aftonsangen@gmail.com`>

Description A central part of using PCA and similar methods are creating and using plots. Visual inspection of these plots is an extremely powerful way of understanding your matrix-form data. Doing this is essential when applying multivariate methods on real-life matrix data. This package provides plots, tailor made for understanding a PCA model on a matrix dataset. We have worked to design plots that supports the mathematical properties of PCA, but still easy to use and understand. The flexibility of the plotting capacity of R has been utilized to create a look and feel that supports the understanding of the PCA model. In addition to that, this package contains a prediction function, using PCA to fill ``holes'' in a matrix. This package does not provide you with the scores and loadings etc in numeric form. This is provided through the underlying package 'mixOmics' which is the basis of this package. In addition, applied PCA on real life data requires flexibility. The original way of calculating a PCA is by using the method SVD, used in e.g. 'stats' package. This is preferred from a precision point of view. But 'PCA4you' is based on the method NIPALS for calculating the PCA model. Hence, the following cases, which are not (or poorly) supported by SVD, are fully supported by 'PCA4you': 1) Non complete matrices with little or a lot of missing values 2) More row than columns, sometimes a huge difference 3) More columns than rows, sometimes a huge difference 4) Complete decomposition is not possible or desired. Sometimes all components can not be calculated for different reasons. Big data, system resources or time can limit the desired number of principal components.

Imports mixOmics

License GPL (>= 2)

Encoding UTF-8

LazyData true

RoxygenNote 6.0.1

NeedsCompilation no

Author Martin Berntsson [aut, cre],
Philip Anton de Saint-Aubain [aut]

Repository CRAN

Date/Publication 2017-10-22 17:02:40 UTC

R topics documented:

`PCA.Filler` *PCA.Filler*

Description

Takes a matrix with NAs and returns a matrix without NAs. Takes data.mat and calculates and returns predicted data.mat. Uses by default a 2 component PCA model. No centering or scaling. Supports NAs and similar. Rownames/colnames are retained. Columns/rows with only NAs (or similar) are removed.

Usage

```
PCA.Filler(data.mat, num.of.pcs = 2)
```

Arguments

`data.mat`	a numeric matrix or data frame. Long, wide and squared matrix supported. NAs supported.
`num.of.pcs`	is an integer specifying the number of components to be used. Default 2.

Author(s)

Berntsson, Martin <martin.berntsson@eon.se> and de Saint-Aubain, Philip Anton <philip-anton.desaint-aubain@knowit.dk>

Examples

```
data.mat <- airquality[, 1:4]
data.mat
PCA.Filler(data.mat)
```

`PCA.Plotter` *PCA.Plotter*

Description

Makes the basic PCA plots on a given matrix. Scores and loadings scatterplots of PC1 vs PC2. Centers and UV-scales the variables, makes a 2 component PCA. Supports NAs but requires variance in all variables. Presents R2 for each component. Uses the package "mixOmics". Rownames/colnames are used as point labels.

Usage

```
PCA.Plotter(data.mat)
```

Arguments

`data.mat`	a numeric matrix or data frame. Long, wide and squared matrix supported. NAs supported. Variance required for all variables.

Author(s)

Berntsson, Martin <martin.berntsson@eon.se> and de Saint-Aubain, Philip Anton <philip-anton.desaint-aubain@knowit.dk>

Examples

```
set.seed(5)
num.of.variables <- 14
num.of.observations <- 34
rn <- rnorm(num.of.variables * num.of.observations)
data.mat <- matrix(rn, nrow = num.of.observations, ncol = num.of.variables)
PCA.Plotter(data.mat)
```

Index