Package ‘LearnPCA’

May 7, 2026

Type Package

Title Functions, Data Sets and Vignettes to Aid in Learning Principal Components Analysis (PCA)

Version 0.3.4

Date 2024-04-25

Description Principal component analysis (PCA) is one of the most widely used data analysis techniques. This package provides a series of vignettes explaining PCA starting from basic concepts. The primary purpose is to serve as a self-study resource for anyone wishing to understand PCA better. A few convenience functions are provided as well.

License GPL-3

URL https://bryanhanson.github.io/LearnPCA/

BugReports https://github.com/bryanhanson/LearnPCA/issues

ByteCompile TRUE

VignetteBuilder knitr

Encoding UTF-8

Imports markdown, shiny, stats, graphics

Suggests ChemoSpec, chemometrics, knitr, tinytest, roxut, rmarkdown, plot3D, ade4, plotrix, latex2exp, plotly, xtable, bookdown,

Depends rpart, class, nnet

RoxygenNote 7.3.1

NeedsCompilation no

Author Bryan A. Hanson [aut, cre] (ORCID: <https://orcid.org/0000-0003-3536-8246>), David T. Harvey [aut]

Maintainer Bryan A. Hanson <hanson@depauw.edu>

Repository CRAN

Date/Publication 2024-04-26 09:00:02 UTC

Contents

LearnPCA-package 2
PCAtoXhat 2
PCsearch 3
XtoPCAtoXhat 4

Index **6**

`LearnPCA-package` *Functions, Data Sets and Vignettes to Aid in Learning Principal Components Analysis (PCA)*

Description

Principal component analysis (PCA) is one of the most widely used data analysis techniques. This package provides a series of vignettes explaining PCA starting from basic concepts. The primary purpose is to serve as a self-study resource for anyone wishing to understand PCA better. A few convenience functions are provided as well.

Author(s)

Bryan A. Hanson and David T. Harvey. Maintainer: Bryan A. Hanson <hanson@depauw.edu>

See Also

Useful links:

- https://bryanhanson.github.io/LearnPCA/
- Report bugs at https://github.com/bryanhanson/LearnPCA/issues

---

`PCAtoXhat` *Use PCA Results to Reconstruct All or Part of the Original Data Set*

Description

This function allows one to reconstruct an approximation (`Xhat`) of the original data using some or all of the principal components, starting from the results of PCA. Inspired by and follows https://stackoverflow.com/a/23603958/633251 very closely. We are grateful for this post by StackOverflow contributor "Marc in the box."

Usage

```
PCAtoXhat(pca, ncomp = NULL)
```

Arguments

pca	An object of class prcomp or princomp (automatically detected). #' The results of data reduction by PCA.
ncomp	Integer. The number of principal components to use in reconstructing the data set. Must be no larger than the number of variables. If not specified, all the components are used and the original data set is reconstructed.

Value

A matrix with the same dimensions as pca$x (the dimensions of the original data set).

Examples

```
# Example data from ?prcomp (see discussion at Stats.StackExchange.com/q/397793)
C <- chol(S <- toeplitz(.9 ^ (0:31)))
set.seed(17)
X <- matrix(rnorm(32000), 1000, 32)
Z <- X %*% C

pcaz <- prcomp(Z)
tst <- PCAtoXhat(pcaz)
all.equal(tst, Z, check.attributes = FALSE)

# Plot to show the effect of increasing ncomp

ntests <- ncol(Z)
rmsd <- rep(NA_real_, ntests)
for (i in 1:ntests) {
ans <- XtoPCAtoXhat(X, i, sd)
del<- ans - X
rmsd[i] <- sqrt(sum(del^2)/length(del)) # RMSD
}
plot(rmsd, type = "b",
  main = "Root Mean Squared Deviation\nReconstructed - Original Data",
  xlab = "No. of Components Retained", ylab = "RMSD")
abline(h = 0.0, col = "pink")
```

PCsearch — *Demonstrate the Search for New Principal Component Axes*

Description

Shiny application to demonstrate the search for the 1st two principal components for a randomly generated set of data.

Usage

```
PCsearch()
```

Details

@return None. A web page opens with the application running.

@author Bryan A. Hanson, David T. Harvey

`XtoPCAtoXhat` *Reduce a Matrix X via PCA and Reconstruct All or Part to Give Xhat*

Description

This function allows one to do "round trip" PCA by reducing a matrix `X` using PCA and then reconstruct an approximation (`Xhat`) using some or all of the principal components. Inspired by https://stats.stackexchange.com/q/229092/26909. We are grateful for this post by StackOverflow contributor Amoeba.

Usage

```
XtoPCAtoXhat(X, ncomp = 3, scale.fun = NULL)
```

Arguments

`X`	A matrix of data, or a structure which can be coerced to a matrix. Samples should be in rows, and variables in columns.
`ncomp`	Integer. The number of principal components to use in reconstructing the data set. Must be no larger than the number of variables.
`scale.fun`	A function to use to scale the data. If `NULL` no scaling will be done.

Value

A matrix with the same dimensions as `X`.

Examples

```
# Example data from ?prcomp (see discussion at Stats.StackExchange.com/q/397793)
C <- chol(S <- toeplitz(.9 ^ (0:31)))
set.seed(17)
X <- matrix(rnorm(32000), 1000, 32)
Z <- X %*% C

tst <- XtoPCAtoXhat(Z)
mean(tst - Z)

# Plot to show the effect of increasing ncomp

ntests <- ncol(Z)
rmsd <- rep(NA_real_, ntests)
for (i in 1:ntests) {
ans <- XtoPCAtoXhat(X, i, sd)
del<- ans - X
```

```
rmsd[i] <- sqrt(sum(del^2)/length(del)) # RMSD
}
plot(rmsd, type = "b",
  main = "Root Mean Squared Deviation\nReconstructed - Original Data",
  xlab = "No. of Components Retained", ylab = "RMSD")
abline(h = 0.0, col = "pink")
```

Index

∗ **multivariate**
LearnPCA-package, 2
∗ **package**
LearnPCA-package, 2

LearnPCA *(*LearnPCA-package*)*, 2
LearnPCA-package, 2

PCAtoXhat, 2
PCsearch, 3

XtoPCAtoXhat, 4