Package 'pcreg'

May 31, 2026

Title Advanced Methods for Principal Component Analysis and Principal Component Regression

Type Package

Version 0.1.1

Author Dr. Pramit Pandit [aut, cre],
Dr. Halagundegowda G R [aut],
Dr. Kamidi Rahul [aut],
Dr. S. Gandhi Doss [aut]

Description Provides a unified framework for principal component analysis (PCA) and principal component regression (PCR), including standard PCA, sparse PCA, robust PCA, and supervised PCA. The package supports automatic selection of the number of components using cumulative variance and elbow methods and integrates PCA with regression modelling through PCR models. It includes tools for PCA suitability assessment using Bartlett's test of sphericity and the Kaiser-Meyer-Olkin (KMO) measure. Visualisation utilities such as scree plots and biplots are provided for interpretation. The methods are designed to handle multicollinearity, outliers, and high-dimensional data, making them suitable for applied statistical modelling and data analysis. The methodology is based on established approaches described in Jolliffe (2002) <doi:10.1007/b98835>, Zou et al. (2006) <doi:10.1111/j.1467-9868.2005.00503.x>, and Hubert et al. (2005) <doi:10.1198/004017004000000563>.

License GPL-3

Encoding UTF-8

RoxygenNote 7.3.3

Imports stats, ggplot2, ggrepel, gridExtra, scales, psych, elasticnet, robustbase, grid

Suggests testthat (>= 3.0.0)

Config/testthat/edition 3

NeedsCompilation no

Maintainer Dr. Pramit Pandit <pramitpandit@gmail.com>

Repository CRAN

Date/Publication 2026-05-31 15:00:08 UTC

Contents

`pca` *Principal Component Analysis (Custom Scree Style)*

Description

Performs PCA with Bartlett test, KMO measure, and visualizes scree plot and biplot (custom style).

Usage

```
pca(data, scale. = TRUE, center = TRUE, verbose = FALSE)
```

Arguments

`data`	A data frame or matrix
`scale.`	Logical; scale variables
`center`	Logical; center variables
`verbose`	Logical; if TRUE, prints Bartlett and KMO results.

Value

List with Bartlett test, KMO, loadings, scores, variance

Examples

```
{
data(mtcars)
Result <- pca(mtcars)
}
```

pcreg *Principal Component Regression (PCR)*

Description

Performs Principal Component Analysis (PCA) on predictors followed by regression using selected principal components. Includes Bartlett's test, KMO measure, scree plot, biplot, and backward model selection.

Usage

```
pcreg(data, y, scale. = TRUE, center = TRUE, verbose = FALSE)
```

Arguments

data	A data frame or matrix
y	Response variable name
scale.	Logical; should variables be scaled?
center	Logical; should variables be centered?
verbose	Logical; print summaries?

Value

A list containing PCA results, regression models, diagnostics, and plots.

Examples

```
data(mtcars)
result <- pcreg(mtcars, y = "mpg")
```

robust_pca *Robust Principal Component Analysis (Robust PCA)*

Description

Performs Robust PCA using MCD estimator with automatic K selection, PCA suitability tests (Bartlett & KMO), scree plot and biplot.

Usage

```
robust_pca(data, K = NULL, threshold = 0.9, verbose = FALSE)
```

Arguments

data	Numeric data frame or matrix
K	Number of components (auto if NULL)
threshold	Variance threshold for K selection
verbose	Logical; if TRUE prints progress messages

Value

List of results

Examples

```
data(iris)
Result <- robust_pca(iris[, 1:4])
```

---

robust_pcreg *Robust Principal Component Regression (Robust PCR)*

Description

Performs Robust PCA using Minimum Covariance Determinant (MCD) followed by regression on robust principal components. Includes PCA suitability tests, summary, loadings, scree plot, biplot, and regression models.

Usage

```
robust_pcreg(data, y, K = NULL, threshold = 0.9, verbose = FALSE)
```

Arguments

data	A data frame or matrix
y	Response variable name (character)
K	Number of components (optional)
threshold	Variance threshold for automatic K selection
verbose	Logical; print detailed output

Value

A list containing Robust PCR outputs

Examples

```
set.seed(123)
dat <- data.frame(
  y = rnorm(100),
  x1 = rnorm(100),
  x2 = rnorm(100),
  x3 = rnorm(100),
  x4 = rnorm(100)
)
Result <- robust_pcreg(dat, y = "y")
```

`sparse_pca` *Sparse Principal Component Analysis (Sparse PCA)*

Description

Performs Sparse PCA using elastic net regularisation with custom scree plot and biplot.

Usage

```
sparse_pca(data, K = NULL, threshold = 0.9, verbose = FALSE)
```

Arguments

`data`	Numeric data frame or matrix
`K`	Number of components (auto if NULL)
`threshold`	Variance threshold for automatic K selection
`verbose`	Logical; prints diagnostic messages

Value

List with PCA results

Examples

```
{
data(mtcars)
Result <- sparse_pca(mtcars)
}
```

sparse_pcreg *Sparse Principal Component Regression (Sparse PCR)*

Description

Performs Sparse PCA followed by regression with automatic selection of number of components. Includes PCA suitability tests, loadings, regression models, scree plot, and biplot.

Usage

```
sparse_pcreg(data, y, K = NULL, threshold = 0.9, verbose = FALSE)
```

Arguments

data	A data frame or matrix
y	Response variable name (character)
K	Number of components (optional)
threshold	Variance threshold for automatic K selection
verbose	Logical; print messages and summaries

Value

A list containing model outputs and PCA results

Examples

```
data(mtcars)
Result <- sparse_pcreg(mtcars, y = "mpg")
```

supervised_pca *Supervised Principal Component Analysis (SPCA)*

Description

True SPCA (Bair et al.): screening + PCA

Usage

```
supervised_pca(data, response, K = NULL, threshold = 0.9, verbose = FALSE)
```

Arguments

data	Numeric data frame or matrix
response	Numeric response vector
K	Number of components (auto if NULL)
threshold	Variance threshold for K selection
verbose	Logical; if TRUE prints progress messages

Value

List of SPCA results

Examples

```
data(mtcars)
Result <- supervised_pca(
  mtcars[, -1],
  mtcars$mpg
)
```

---

supervised_pcreg *Supervised Principal Component Regression*

Description

Performs supervised feature selection followed by Principal Component Regression (PCR). Includes Bartlett's test, KMO measure, scree plot, biplot, and regression models.

Usage

```
supervised_pcreg(data, response, K = NULL, threshold = 0.9, verbose = FALSE)
```

Arguments

data	A data frame or matrix
response	Response variable name
K	Number of principal components
threshold	Cumulative variance threshold for automatic component selection
verbose	Logical; print summaries?

Value

A list containing PCA results, regression models, diagnostics, and plots.

Examples

```
data(mtcars)
result <- supervised_pcreg(
  mtcars,
  response = "mpg"
)
```

Index