Package 'SparseBiplots'

July 11, 2025

Type Package

Title 'HJ-Biplot' using Different Ways of Penalization Plotting with 'ggplot2'

Version 4.1.1

Maintainer Mitzi Isabel Cubilla-Montilla <mitzi@usal.es>

Description The 'HJ-Biplot' is a multivariate method that represents high-dimensional data in a low-dimensional subspace, capturing most of the information's variability in just a few dimensions. This package implements three new regularized versions of the HJ-Biplot: Ridge, LASSO, and Elastic Net. These versions introduce restrictions that shrink or zero-out variable weights to improve interpretability based on regularization theory. All methods provide graphical representations using 'ggplot2'.

License GPL (>= 3)

Encoding UTF-8

Depends R (>= 3.3.0), ggplot2

Imports ggrepel, stats, sparsepca

Suggests testthat

URL https://github.com/mitzicubillamontilla/SparseBiplots

BugReports https://github.com/mitzicubillamontilla/SparseBiplots/issues

RoxygenNote 7.3.2

NeedsCompilation no

Author Mitzi Isabel Cubilla-Montilla [aut, cre],
Carlos Alfredo Torres-Cubilla [aut],
Maria Purificación Galindo Villardón [aut],
Ana Belen Nieto-Librero [aut]

Repository CRAN

Date/Publication 2025-07-11 04:40:02 UTC

Contents

ElasticNet_HJBiplot   *Elastic Net HJ Biplot*

Description

This function is a generalization of the Ridge regularization method and the LASSO penalty. Realizes the representation of the SPARSE HJ Biplot through a combination of LASSO and Ridge, on the data matrix. This means that with this function you can eliminate weak variables completely as with the LASSO regularization or contract them to zero as in Ridge.

Usage

```
ElasticNet_HJBiplot(X, Lambda = 1e-04, Alpha = 1e-04, Transform.Data = 'scale')
```

Arguments

X	array_like; A data frame with the information to be analyzed
Lambda	float; Tuning parameter of the LASSO penalty. Higher values lead to sparser components.
Alpha	float; Tuning parameter of the Ridge shrinkage
Transform.Data	character; A value indicating whether the columns of X (variables) should be centered or scaled. Options are: "center" that removes the columns means and "scale" that removes the columns means and divide by its standard deviation. Default is "scale".

Details

Algorithm used to perform automatic selection of variables and continuous contraction simultaneously. With this method, the model obtained is simpler and more interpretable. It is a particularly useful method when the number of variables is much greater than the number of observations.

Value

`ElasticNet_HJBiplot` returns a list containing the following components:

`loadings`	array_like; penalized loadings, the loadings of the sparse principal components.
`n_ceros`	array_like; number of loadings equal to cero in each component.
`coord_ind`	array_like; matrix with the coordinates of individuals.
`coord_var`	array_like; matrix with the coordinates of variables.
`eigenvalues`	array_like; vector with the eigenvalues penalized.
`explvar`	array_like; an vector containing the proportion of variance explained by the first 1, 2,.,k sparse principal components obtained.

Author(s)

Mitzi Cubilla-Montilla, Carlos Torres-Cubilla, Ana Belen Nieto Librero and Purificacion Galindo Villardon

References

- Galindo, M. P. (1986). Una alternativa de representacion simultanea: HJ-Biplot. Questiio, 10(1), 13-23.

- Erichson, N. B., Zheng, P., Manohar, K., Brunton, S. L., Kutz, J. N., & Aravkin, A. Y. (2018). Sparse principal component analysis via variable projection. arXiv preprint arXiv:1804.00341.

- Zou, H., & Hastie, T. (2005). Regularization and variable selection via the elastic net. Journal of the royal statistical society: series B (statistical methodology), 67(2), 301-320.

See Also

[elasticnet]{spca}, `Plot_Biplot`

Examples

```
ElasticNet_HJBiplot(mtcars, Lambda = 0.2, Alpha = 0.1)
```

HJBiplot   *HJ Biplot*

Description

This function performs the representation of HJ Biplot (Galindo, 1986).

Usage

```
HJBiplot (X, Transform.Data = 'scale')
```

Arguments

X   array_like;
                        A data frame which provides the data to be analyzed. All the variables must be
                        numeric.

Transform.Data character;
                        A value indicating whether the columns of X (variables) should be centered or
                        scaled. Options are: "center" that removes the columns means and "scale" that
                        removes the columns means and divide by its standard deviation. Default is
                        "scale".

Details

Algorithm used to construct the HJ Biplot. The Biplot is obtained as result of the configuration of
markers for individuals and markers for variables in a reference system defined by the factorial axes
resulting from the Decomposition in Singular Values (DVS).

Value

HJBiplot returns a list containing the following components:

eigenvalues   array_like;
                vector with the eigenvalues.

explvar   array_like;
                an vector containing the proportion of variance explained by the first 1, 2,.,k
                principal components obtained.

loadings   array_like;
                the loadings of the principal components.

coord_ind   array_like;
                matrix with the coordinates of individuals.

coord_var   array_like;
                matrix with the coordinates of variables.

Author(s)

Mitzi Cubilla-Montilla, Carlos Torres-Cubilla, Ana Belen Nieto Librero and Purificacion Galindo Villardon

References

- Gabriel, K. R. (1971). The Biplot graphic display of matrices with applications to principal components analysis. Biometrika, 58(3), 453-467.
- Galindo, M. P. (1986). Una alternativa de representacion simultanea: HJ-Biplot. Questiio, 10(1), 13-23.

See Also

[Plot_Biplot]

Examples

```
HJBiplot(mtcars)
```

LASSO_HJBiplot   *LASSO HJ Biplot*

Description

This function performs the representation of the SPARSE HJ Biplot applying the LASSO regularization, on the original data matrix, implementing the norm L1.

Usage

```
LASSO_HJBiplot(X, Lambda, Transform.Data = 'scale', Operator = 'Hard-Thresholding')
```

Arguments

X	array_like; A data frame which provides the data to be analyzed. All the variables must be numeric.
Lambda	float; Tuning parameter for the LASSO penalty
Transform.Data	character; A value indicating whether the columns of X (variables) should be centered or scaled. Options are: "center" that removes the columns means and "scale" that removes the columns means and divide by its standard deviation. Default is "scale".
Operator	character; The operator used to solve the norm L1. Allowed values are "Soft-Thresholding" and "Hard-Thresholding".

Details

Algorithm that performs a procedure of contraction and selection of variables. LASSO imposes a penalty that causes the charges of some components to be reduced to zero. By producing zero loadings for some components and not zero for others, the Lasso technique performs selection of variables. As the value of the penalty approaches one, the loadings approach zero.

Value

LASSO_HJBiplot returns a list containing the following components:

loadings	array_like; penalized loadings, the loadings of the sparse principal components.
n_ceros	array_like; number of loadings equal to cero in each component.
coord_ind	array_like; matrix with the coordinates of individuals.
coord_var	array_like; matrix with the coordinates of variables.
eigenvalues	array_like; vector with the eigenvalues penalized.
explvar	array_like; an vector containing the proportion of variance explained by the first 1, 2,.,k sparse principal components obtained.

Author(s)

Mitzi Cubilla-Montilla, Carlos Torres-Cubilla, Ana Belen Nieto Librero and Purificacion Galindo Villardon

References

- Galindo, M. P. (1986). Una alternativa de representacion simultanea: HJ-Biplot. Questiio, 10(1), 13-23.

- Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. Journal of the Royal Statistical Society: Series B (Methodological), 58(1), 267-288.

- Tibshirani, R. (2011). Regression shrinkage and selection via the lasso: a retrospective. Journal of the Royal Statistical Society: Series B (Statistical Methodology), 73(3), 273-282.

See Also

Plot_Biplot

Examples

```
LASSO_HJBiplot(mtcars, Lambda = 0.2, Operator = 'Hard-Thresholding')
```

Plot_Biplot   *Plotting Biplot*

Description

`Plot_Biplot` initializes a ggplot2-based visualization of the caracteristics presented in the data analized by the Biplot selected.

Usage

```
Plot_Biplot(X, axis = c(1,2), hide = "none",
 labels = "auto", ind.shape = 19,
 ind.color = "red", ind.size = 2,
 ind.label = FALSE, ind.label.size = 4,
 var.color = "black", var.size = 0.5,
 var.arrow.type = "open", var.arrow.angle = 30, var.arrow.length = 0.25,
 var.label = TRUE, var.label.size = 4, var.label.angle = FALSE)
```

Arguments

X	List containing the output of one of the functions of the package.
axis	Vector with lenght 2 which contains the axis ploted in x and y axis.
hide	Vector specifying the elements to be hidden on the plot. Default value is "none". Other allowed values are "ind" and "var".
labels	It indicates the label for points. If it is "auto" the labels are the row names of the coordinates of individuals. If it isn't auto it would be a vector containing the labels.
ind.shape	Points shape. It can be a number to indicate the shape of all the points or a factor to indicate different shapes.
ind.color	Points colors. It can be a character indicating the color of all the points or a factor to use different colors.
ind.size	Size of points.
ind.label	Logical value, if it is TRUE it prints the name for each row of X. If it is FALSE (default) does not print the names.
ind.label.size	Numeric value indicating the size of the labels of points.
var.color	Character indicating the color of the arrows.
var.size	Size of arrow.
var.arrow.type	One of "open" or "closed" indicating whether the arrow head should be a closed triangle.
var.arrow.angle	
	The angle of the arrow head in degrees (smaller numbers produce narrower, pointier arrows). Essentially describes the width of the arrow head.

```
var.arrow.length
```
> Length of the arrow head (in cm).

```
var.label
```
> Logical value, if it is TRUE (default) it prints the name for each column of X. If it is FALSE does not print the names.

```
var.label.size  Numeric value indicating the size of the labels of variables.
var.label.angle
```
> Logical value, if it it TRUE (default) it print the vector names with orentation of the angle of the vector. If it is FALSE the angle of all tags is 0.

Value

Return a **ggplot2** object.

Author(s)

Mitzi Cubilla-Montilla, Carlos Torres-Cubilla, Ana Belen Nieto Librero and Purificacion Galindo Villardon

See Also

[HJBiplot](#), [Ridge_HJBiplot](#), [ElasticNet_HJBiplot](#)

Examples

```
hj.biplot <- HJBiplot(mtcars)
Plot_Biplot(hj.biplot, ind.label = TRUE)
```

Ridge_HJBiplot   *Ridge HJ Biplot*

Description

This function performs the representation of the HJ Biplot applying the Ridge regularization, on the original data matrix, implementing the norm L2.

Usage

```
Ridge_HJBiplot (X, Lambda, Transform.Data = 'scale')
```

Arguments

```
X
```
> array_like;
> A data frame which provides the data to be analyzed. All the variables must be numeric.

```
Lambda
```
> float;
> Tuning parameter for the Ridge penalty

Transform.Data character;

> A value indicating whether the columns of X (variables) should be centered or scaled. Options are: "center" that removes the columns means and "scale" that removes the columns means and divide by its standard deviation. Default is "scale".

Details

Algorithm used to contract the loads of the main components towards zero, but without achieving the nullity of any. If the penalty parameter is less than or equal to 1e-4 the result is like Galindo's HJ Biplot (1986).

Value

Ridge_HJBiplot returns a list containing the following components:

eigenvalues array_like;
> vector with the eigenvalues penalized.

explvar array_like;
> an vector containing the proportion of variance explained by the first 1, 2,,,k sparse principal components obtained.

loadings array_like;
> penalized loadings, the loadings of the sparse principal components.

coord_ind array_like;
> matrix with the coordinates of individuals.

coord_var array_like;
> matrix with the coordinates of variables.

Author(s)

Mitzi Cubilla-Montilla, Carlos Torres-Cubilla, Ana Belen Nieto Librero and Purificacion Galindo Villardon

References

- Galindo, M. P. (1986). Una alternativa de representacion simultanea: HJ-Biplot. Questiio, 10(1), 13-23.

- Hoerl, A. E., & Kennard, R. W. (1970). Ridge regression: Biased estimation for nonorthogonal problems. Technometrics, 12(1), 55-67.

- Zou, H., Hastie, T., & Tibshirani, R. (2006). Sparse principal component analysis. Journal of computational and graphical statistics, 15(2), 265-286.

See Also

Plot_Biplot

Examples

```
Ridge_HJBiplot(mtcars, Lambda = 0.2)
```

Index