Package ‘areabiplot’

October 12, 2022

Title Area Biplot

Version 1.0.0

Description Considering an (n x m) data matrix X, this package is based on the method proposed by Gower, Groener, and Velden (2010) <doi:10.1198/jcgs.2010.07134>, and utilize the resulting matrices from the extended version of the NIPALS decomposition to determine n triangles whose areas are used to visually estimate the elements of a specific column of X. After a 90-degree rotation of the sample points, the triangles are drawn regarding the following points: 1.the origin of the axes; 2.the sample points; 3. the vector endpoint representing some variable.

License MIT + file LICENSE

Encoding UTF-8

LazyData true

RoxygenNote 7.1.1

Suggests testthat

Imports grDevices, graphics, nipals

NeedsCompilation no

Author Alberto Silva [aut, cre] (<https://orcid.org/0000-0002-3496-6802>), Adelaide Freitas [aut] (<https://orcid.org/0000-0002-4685-1615>)

Maintainer Alberto Silva <albertos@ua.pt>

Repository CRAN

Date/Publication 2021-03-10 19:00:02 UTC

R topics documented:

`areabiplot` *Area Biplot*

Description

Consider an (n x m) centered data matrix X and let $rank(X) = r$. Alternatively to the ordinary NIPALS decomposition of X, where $X = TP'$, this package uses the resulting matrices from the extended version of the NIPALS decomposition ($X = GHP'$) to determine n triangles whose areas are used to visually estimate the n elements of a specific column of X (a variable of interest). After a 90-degree rotation of the sample points, the triangles are drawn regarding the following points:

1. the origin of the axes.
2. the sample points.
3. the vector endpoint representing the selected variable.

Just keep in mind that The extended NIPALS decomposition, $X = GHP'$, is equivalent to the SVD decomposition, $X = UDV'$, being that:

1. G is the matrix containing in its columns the normalized score vectors of X, i.e., the normalized columns of T. If t is the i-th score vector of the matrix T, then the i-th column of G will be $g = t/||t||$, which will correspond to the i-th left singular vector u.
2. If t is the i-th column of T, then $||t|| = \sqrt{(t't)}$ gives the i-th singular value of X. In addition, H is the diagonal matrix containing these singular values in decreasing order, i.e., $H = D$.
3. P is the loadings matrix, which is equivalent to the V matrix that contains the right singular vectors of X.

Usage

```
areabiplot(
  L,
  S,
  R,
  ord.row,
  mode = NULL,
  tri.rgb = NULL,
  bg.col = NULL,
  plot.title = NULL,
  plot.title.col = NULL,
  plot.title.font = NULL,
  plot.title.cex = NULL,
  plot.sub = NULL,
  plot.sub.col = NULL,
  plot.sub.font = NULL,
  plot.sub.cex = NULL,
  plot.cex = NULL,
  plot.col = NULL,
```

```
  plot.pch = NULL,
  plot.xlab = NULL,
  plot.ylab = NULL,
  plot.xlim = NULL,
  plot.ylim = NULL,
  points.lab = NULL,
  var.lab = NULL,
  text.col.var = NULL,
  text.cex = NULL,
  text.font = NULL,
  text.col = NULL,
  text.pos = NULL,
  axis.col = NULL,
  axis.cex = NULL,
  axis.font = NULL,
  axis.asp = NULL,
  arrow.lwd = NULL,
  arrow.len = NULL,
  arrow.col = NULL
)
```

Arguments

`L` A (n x 2) matrix containing normalized score vectors g (or left singular vectors u).

`S` An appropriate (2 x 2) diagonal matrix containing the corresponding singular values in decreasing order.

`R` A (m x 2) matrix containing the corresponding loading vectors (or right singular vectors).

`ord.row` The row of R used as the base of the triangle, e.g., if 1 is provided, then the first row of R will be taken.

`mode` a string providing the way the singular values will be allocated. The default is "SS", i.e., the similar spread proposed by Gower et al.. Alternatively, one can choose the "HJ" method (see more in Details).

`tri.rgb` The hexadecimal color and alpha transparency code for the triangle. The default is #19FF811A (green and 90% of transparency).

`bg.col` A string providing the color of the background. The default is #001F3D (blue).

`plot.title` A string providing the main title. The default is NONE.

`plot.title.col` A string specifying the color of the main title text. The default is "FFFFFF" (white).

`plot.title.font` An integer providing the style of the main title text. The default is 1 (normal text).

`plot.title.cex` A number indicating the amount by which the main title text should be scaled relative to the default. 1 = default, 1.5 is 50% larger, and so on.

`plot.sub` A string providing a sub-title. The default is NONE.

`plot.sub.col`	A string specifying the color of the sub-title text. The default is "FFFFFF" (white).
`plot.sub.font`	An integer providing the style of the main title text. The default is 1 (normal text).
`plot.sub.cex`	A number indicating the amount by which the sun-title text should be scaled relative to the default. 1 = default, 1.5 is 50% larger, and so on.
`plot.cex`	A number indicating the expansion or contraction factor used to specify the point size. The default is 0.6 (40% smaller).
`plot.col`	A string specifying the color of the points. The default is "FFFFFF" (white).
`plot.pch`	An integer specifying the shape of the points. The default is 21 (circle)
`plot.xlab`	A string specifying a label to the horizontal axis. The default is NONE.
`plot.ylab`	A string specifying a label to the vertical axis. The default is NONE.
`plot.xlim`	The limits for the x axis.
`plot.ylim`	The limits for the y axis.
`points.lab`	A vector of characters containing the names of the data matrix rows.
`var.lab`	A string providing the variable name used as triangle base.
`text.col.var`	A string specifying the color of the variable label text. The default is "FFFFFF" (white).
`text.cex`	A number indicating the expansion or contraction factor used to specify the point labels. The default is 0.5.
`text.font`	An integer providing the style of the point labels. The default is 2 (bold).
`text.col`	A string specifying the color of the point labels text. The default is "FFFFFF" (white).
`text.pos`	An integer providing the position of the point labels. The default is 3 (above).
`axis.col`	A string specifying the color of the axis. The default is #FFFFFF (white).
`axis.cex`	A number indicating the expansion or contraction factor used to specify the tick label. The default is 0.7
`axis.font`	An integer providing the style of the tick label. The default is 1 (normal text).
`axis.asp`	A number specifying the aspect ratio of the axes. The default is 1.
`arrow.lwd`	A number specifying the line width of the arrow. The default is 1.
`arrow.len`	The length of the edges of the arrow head (in inches). The default is 0.1.
`arrow.col`	A string specifying the color of the arrow. The default is "FFFFFF" (white).

Details

1. If the variables (the columns of X) are measured in different units or their variability differs considerably, one could perform a variance scaling to get better visual results on the graph (see Examples). In this case, the percentage of variance explained by the first principal components might decrease.
2. The "HJ" mode is reserved for an application under implementation.

Value

An area biplot is produced on the current graphics device.

Author(s)

Alberto Silva albertos@ua.pt, Adelaide Freitas adelaide@ua.pt

References

J.C. Gower, P.J.F. Groenen, M. van de Velden (2010). Area Biplots. Journal of Computational and Graphical Statistics, v.19 (1), pp. 46-61. doi: 10.1198/jcgs.2010.07134

Examples

```
library(nipals)
data(uscrime)
Y = uscrime[, -1]

# first case: scale is false
nip = nipals(Y, ncomp = 2, center = TRUE, scale = FALSE, force.na = TRUE)
L = nip$scores
R = nip$loadings
S = diag(nip$eig[1:2])
areabiplot(L, S, R, 5, points.lab = c(uscrime[, 1]),var.lab= "burglary")

# second case: scale is true
nip = nipals(Y, ncomp = 2, center = TRUE, scale = TRUE, force.na = TRUE)
L = nip$scores
R = nip$loadings
S = diag(nip$eig[1:2])
areabiplot(L, S, R, 4, points.lab = c(uscrime[, 1]),var.lab= "assault")
```

Index