

Package ‘calibrate’

June 19, 2020

Version 1.7.7

Date 2020-06-18

Title Calibration of Scatterplot and Biplot Axes

Author Jan Graffelman <jan.graffelman@upc.edu>

Maintainer Jan Graffelman <jan.graffelman@upc.edu>

Depends R (>= 3.5.0), MASS

Description Package for drawing calibrated scales with tick marks on (non-orthogonal) variable vectors in scatterplots and biplots. Also provides some functions for biplot creation and for multivariate analysis such as principal coordinate analysis.

License GPL-2

NeedsCompilation no

Repository CRAN

Date/Publication 2020-06-19 05:40:14 UTC

R topics documented:

bplot	2
calibrate	3
calves	6
canocor	6
circle	8
dlines	9
goblets	10
heads	10
linnerud	11
ones	11
origin	12
PrinCoor	13
rad2degree	14
rda	15
shiftvector	16
spaindist	17
storks	18
textxy	18

bplot

*General function for making biplots***Description**

Function `bplot` creates biplots on the basis matrices of row and column markers.

Usage

```
bplot(Fr,G,rowlab=rownames(Fr),collab=rownames(G),qlt=rep(1,nrow(Fr)),
refaxis=TRUE,ahead=T,xl=NULL,yl=NULL,frame=F,qltlim=0,rowch=19,
colch=19,qltvar=NULL,rowcolor="red",colcolor="blue",rowmark=TRUE,
colmark=TRUE,rowarrow=FALSE,colarrow=TRUE,markrowlab=TRUE,
markcollab=TRUE,xlab="",ylab="",cex.rowlab=1,cex.rowdot=0.75,
cex.collab=1,cex.coldot=0.75,cex.axis=0.75,lwd=1,arrowangle=10,...)
```

Arguments

<code>Fr</code>	matrix with coordinates of the row markers.
<code>G</code>	matrix with coordinates of the column markers.
<code>rowlab</code>	vector with labels for the rows.
<code>collab</code>	vector with labels for the columns.
<code>qlt</code>	goodness of fit of the rows.
<code>refaxis</code>	draw coordinate system <code>refaxis=TRUE</code> or not.
<code>ahead</code>	put a head on the vectors <code>ahead=TRUE</code> or not.
<code>xl</code>	limits for the x-axis.
<code>yl</code>	limits for the y-axis.
<code>frame</code>	draw a box around the plot <code>frame=TRUE</code> or not.
<code>qltlim</code>	draw only the vectors with a goodness of fit larger than <code>qltlim</code> .
<code>rowch</code>	character used for the row markers.
<code>colch</code>	character used for the column markers.
<code>qltvar</code>	vector with the goodness of fit of each variable.
<code>rowcolor</code>	colour used for the row markers.
<code>colcolor</code>	colour used for the column markers.
<code>rowmark</code>	show row markers (<code>rowmark=TRUE</code>) or not.
<code>colmark</code>	show column markers (<code>colmark=TRUE</code>) or not.
<code>rowarrow</code>	draw vectors from the origin to the row markers (<code>rowarrow=TRUE</code>) or not.
<code>colarrow</code>	draw vectors from the origin to the column markers (<code>colarrow=TRUE</code>) or not.
<code>markrowlab</code>	depict row marker labels (<code>rowlab=TRUE</code>) or not.

markcollab	depict column marker labels (collab=TRUE) or not.
xlab	a label for the x-axis.
ylab	a label for the y-axis.
cex.rowlab	expansion factor for the row labels.
cex.rowdot	expansion factor for the row markers.
cex.collab	expansion factor for the column labels.
cex.coldot	expansion factor for the column markers.
cex.axis	expansion factor for the axis.
lwd	line width for biplot vectors.
arrowangle	angle for the edges of the arrowhead.
...	extra arguments for plot.

Value

None. The function produces a graphic.

Author(s)

Jan Graffelman (jan.graffelman@upc.edu)

Examples

```
set.seed(123)
X <- matrix(runif(40),byrow=TRUE,ncol=4)
colnames(X) <- paste("X",1:ncol(X),sep="")
out.pca <- princomp(X,cor=TRUE)
Fp <- out.pca$scores
Gs <- as.matrix(unclass(out.pca$loadings))
bplot(Fp,Gs,colch=NA)
```

calibrate

Calibration of Biplot and Scatterplot Axis

Description

Routine for the calibration of any axis (variable vector) in a biplot or a scatterplot

Usage

```
calibrate(g,y,tm,Fr,tmlab=tm,tl=0.05,dt=TRUE,dp=FALSE,
  lm=TRUE,verb=TRUE,axislab="",reverse=FALSE,
  alpha=NULL,labpos=1,weights=diag(rep(1,length(y))),
  axiscol="blue",cex.axislab=0.75,graphics=TRUE,where=3,
  laboffset=c(0,0),m=matrix(c(0,0),nrow=1),markerpos=3,
  showlabel=TRUE,lwd=1,shiftvec=c(0,0),shiftdir="none",shiftfactor=1.05)
```

Arguments

<code>g</code>	the vector to be calibrated (2 x 1).
<code>y</code>	the data vector corresponding to <code>g</code> , appropriately centred and/or standardized.
<code>tm</code>	the vector of tick marks, appropriately centred and/or scaled.
<code>Fr</code>	the coordinates of the rows markers in the biplot.
<code>tmlab</code>	a list or vector of tick mark labels.
<code>t1</code>	the tick length. By default, the tick markers have length 0.05.
<code>dt</code>	draw ticks. By default, ticks markers are drawn. Set <code>dt=F</code> in order to compute calibration results without actually drawing the calibrated scale.
<code>dp</code>	drop perpendiculars. With <code>dp=T</code> perpendicular lines will be drawn from the row markers specified by <code>Fr</code> onto the calibrated axis. This is a graphical aid to read off the values in the corresponding scale.
<code>lm</code>	label markers. By default, all tick marks are labelled. Setting <code>lm=F</code> turns off the labelling of the tick marks. This allows for creating tick marks without labels. It is particularly useful for creating finer scales of tickmarks without labels.
<code>verb</code>	verbose parameter (<code>F</code> =be quiet, <code>T</code> =show results).
<code>axislab</code>	a label for the calibrated axis.
<code>reverse</code>	puts the tick marks and tick mark labels on the other side of the axis.
<code>alpha</code>	a value for the calibration factor. This parameter should only be specified if a calibration is required that is different from the one that is optimal for data recovery.
<code>labpos</code>	position of the label for the calibrated axis (1,2,3 or 4).
<code>laboffset</code>	offset vector for the axis label. If specified, shifts the label by the specified amounts with respect to the current position.
<code>weights</code>	a matrix of weights (optional).
<code>axiscol</code>	color of the calibrated axis.
<code>cex.axislab</code>	character expansion factor for axis label and tick mark labels.
<code>graphics</code>	do graphics or not (<code>F</code> =no graphical output, <code>T</code> =draws calibrated scale).
<code>where</code>	label placement (1=beginning,2=middle,3=end).
<code>m</code>	vector of means.
<code>markerpos</code>	position specifier for the tick mark labels (1,2,3 or 4).
<code>showlabel</code>	show axis label in graph (<code>T</code>) or not (<code>F</code>).
<code>lwd</code>	line width for the calibrated axis
<code>shiftvec</code>	a shift vector for the calibrated axis ((0,0) by default)
<code>shiftdir</code>	indicates in which direction the axis should be shifted ("left", "right" or "none"). This direction is w.r.t. vector <code>g</code>
<code>shiftfactor</code>	scalar by which the shift vector is stretched (or shrunken). By default, the length of the shift vector is stretched by 5 percent (<code>shiftfactor = 1.05</code>)

Details

This program calibrates variable vectors in biplots and scatterplots, by drawing tick marks along a given the vector and labelling the tick marks with specified values. The optimal calibration is found by (generalized) least squares. Non-optimal calibrations are possible by specifying a calibration factor (alpha).

Value

Returns a list with calibration results

useralpha	calibration factor specified by the user
optalpha	optimal calibration factor
lengthoneunit	length in the plot of one unit in the scale of the calibrated variable
gof	goodness of fit (as in regression)
gos	goodness of scale
M	coordinates of the tick markers
ang	angle in degrees of the biplot axis with the positive x-axis
shiftvec	the supplied or computed shift vector
yt	fitted values for the variable according to the calibration
e	errors according to the calibration
Fpr	coordinates of the projections of the row markers onto the calibrated axis
Mn	coordinates of the tick marker end points

Author(s)

Jan Graffelman <jan.graffelman@upc.edu>

References

Gower, J.C. and Hand, D.J., (1996) Biplots. Chapman & Hall, London

Graffelman, J. and van Eeuwijk, F.A. (2005) Calibration of multivariate scatter plots for exploratory analysis of relations within and between sets of variables in genomic research *Biometrical Journal*, 47(6) pp. 863-879.

Graffelman, J. (2006) A guide to biplot calibration.

See Also

[biplot](#)

Examples

```
x <- rnorm(20,1)
y <- rnorm(20,1)
x <- x - mean(x)
y <- y - mean(y)
z <- x + y
b <- c(1,1)
plot(x,y,asp=1,pch=19)
tm<-seq(-2,2,by=0.5)
Calibrate.z <- calibrate(b,z,tm,cbind(x,y),axislab="Z",graphics=TRUE)
```

calves	<i>Delivery of Dutch Calves</i>
--------	---------------------------------

Description

This data set gives a cross classification of 7275 calves born in the late nineties according to type of production and type of delivery.

Usage

```
data(calves)
```

Format

A data frame containing a contingency table of 7275 observations.

Source

Holland Genetics. <http://www.hg.nl>

References

Graffelman, J. (2005) *A guide to scatterplot and biplot calibration*.

canocor	<i>Canonical correlation analysis</i>
---------	---------------------------------------

Description

canocor performs canonical correlation analysis on the basis of the standardized variables and stores extensive output in a list object.

Usage

```
canocor(X, Y)
```

Arguments

X	a matrix containing the X variables
Y	a matrix containing the Y variables

Details

canocor computes the solution by a singular value decomposition of the transformed between set correlation matrix.

Value

Returns a list with the following results

ccor	the canonical correlations
A	canonical weights of the x variables
B	canonical weights of the y variables
U	canonical x variates
V	canonical y variates
Fs	biplot markers for x variables (standard coordinates)
Gs	biplot markers for y variables (standard coordinates)
Fp	biplot markers for x variables (principal coordinates)
Gp	biplot markers for y variables (principal coordinates)
fitRxy	goodness of fit of the between-set correlation matrix
fitXs	adequacy coefficients of x variables
fitXp	redundancy coefficients of x variables
fitYs	adequacy coefficients of y variables
fitYp	redundancy coefficients of y variables

Author(s)

Jan Graffelman <jan.graffelman@upc.edu>

References

- Hotelling, H. (1935) The most predictable criterion. *Journal of Educational Psychology* (26) pp. 139-142.
- Hotelling, H. (1936) Relations between two sets of variates. *Biometrika* (28) pp. 321-377.
- Johnson, R. A. and Wichern, D. W. (2002) *Applied Multivariate Statistical Analysis*. New Jersey: Prentice Hall.

See Also

[cancor](#)

Examples

```
set.seed(123)
X <- matrix(runif(75),ncol=3)
Y <- matrix(runif(75),ncol=3)
cca.results <- canocor(X,Y)
```

circle

Draw a circle

Description

circle draws a circle in an existing plot.

Usage

```
circle(radius,origin)
```

Arguments

radius	the radius of the circle
origin	the origin of the circle

Value

NULL

Author(s)

Jan Graffelman <jan.graffelman@upc.edu>

Examples

```
set.seed(123)
X <- matrix(rnorm(20),ncol=2)
plot(X[,1],X[,2])
circle(1,c(0,0))
```

dlines	<i>Connect two sets of points by lines</i>
--------	--

Description

dlines connects two sets of points by lines in a rowwise manner.

Usage

```
dlines(SetA, SetB, lin = "dotted")
```

Arguments

SetA	matrix with the first set of points
SetB	matrix with teh second set of points
lin	linestyle for the connecting lines

Value

NULL

Author(s)

Jan Graffelman (jan.graffelman@upc.edu)

See Also

[lines](#)

Examples

```
X <- matrix(runif(20),ncol=2)
Y <- matrix(runif(20),ncol=2)
plot(rbind(X,Y))
text(X[,1],X[,2],paste("X",1:10,sep=""))
text(Y[,1],Y[,2],paste("Y",1:10,sep=""))
dlines(X,Y)
```

goblets

Size measurements of archeological goblets

Description

This data set gives 6 different size measurements of 25 goblets

Usage

```
data(goblets)
```

Format

A data frame containing 25 observations.

Source

Manly, 1989

References

Manly, B. F. J. (1989) *Multivariate statistical methods: a primer*. London: Chapman and Hall, London

heads

Dimensions of heads of first and second sons for 25 families

Description

Variables X1 and X2 are the head length and head breadth of the first son and Y1 and Y2 are the same variables for the second son.

Usage

```
data(heads)
```

Format

A data frame containing 25 observations.

Source

Mardia, 1979, p. 121

References

- Frets, G. P. (1921) *Heredity of head form in man*, *Genetica* 3, pp. 193-384.
- Mardia, K. V. and Kent, J. T. and Bibby, J. M. (1979) *Multivariate Analysis*. Academic Press London.
- Anderson, T. W. (1984) *An Introduction to Multivariate Statistical Analysis*. New York: John Wiley, Second edition.

linnerud

Linnerud's exercise and body measurements

Description

The data set consist of 3 exercise variables (Tractions a la barre fixe, Flexions, Sauts) and 3 body measurements (Poids, Tour de taille, Pouls) of 20 individuals.

Usage

```
data(linnerud)
```

Format

A data frame containing 20 observations.

Source

Tenenhaus, 1998, table 1, page 15

References

- Tenenhaus, M. (1998) *La Regression PLS*. Paris: Editions Technip.

ones

Generates a matrix of ones

Description

ones generates a matrix of ones.

Usage

```
ones(n, p = n)
```

Arguments

n number of rows
p number of columns

Details

if only `n` is specified, the resulting matrix will be square.

Value

a matrix filled with ones.

Author(s)

Jan Graffelman (jan.graffelman@upc.edu)

See Also

[matrix](#)

Examples

```
Id <- ones(3)
print(Id)
```

origin

Origin

Description

Draws coordinate axes in a plot.

Usage

```
origin(m=c(0,0), ...)
```

Arguments

<code>m</code>	the coordinates of the means (2 x 1).
<code>...</code>	other arguments passed on to the <code>lines</code> function

Author(s)

Jan Graffelman (jan.graffelman@upc.edu)

See Also

[lines](#)

Examples

```
X <- matrix(runif(40), ncol=2)
plot(X[,1], X[,2])
origin(m=c(mean(X[,1]), mean(X[,2])))
```

PrinCoor

*Function for Principal Coordinate Analysis***Description**

Function PrinCoor implements Principal Coordinate Analysis, also known as classical metric multidimensional scaling or classical scaling. In comparison with other software, it offers refined statistics for goodness-of-fit at the level of individual observations and pairs of observations.

Usage

```
PrinCoor(Dis, eps = 1e-10)
```

Arguments

Dis	A distance matrix or dissimilarity matrix
eps	A tolerance criterion for deciding if eigenvalues are zero or not

Details

Calculations are based on the spectral decomposition of the scalar product matrix B, derived from the distance matrix.

Value

X	The coordinates of the the solution
la	The eigenvalues of the solution
B	The scalar product matrix
standard.decom	Standard overall goodness-of-fit table using all eigenvalues
positive.decom	Overall goodness-of-fit table using only positive eigenvalues
absolute.decom	Overall goodness-of-fit table using absolute values of eigenvalues
squared.decom	Overall goodness-of-fit table using squared eigenvalues
RowStats	Detailed goodness-of-fit statistics for each row
PairStats	Detailed goodness-of-fit statistics for each pair

Author(s)

Jan Graffelman <jan.graffelman@upc.edu>

References

Graffelman, J. (2019) Goodness-of-fit filtering in classical metric multidimensional scaling with large datasets. <doi: 10.1101/708339>

Graffelman, J. and van Eeuwijk, F.A. (2005) Calibration of multivariate scatter plots for exploratory analysis of relations within and between sets of variables in genomic research *Biometrical Journal*, 47(6) pp. 863-879.

See Also[princomp](#)**Examples**

```
data(spaindist)
results <- PrinCoor(as.matrix(spaindist))
```

`rad2degree`*Convert radians to degrees.*

Description

`rad2degree` converts radians to degrees.

Usage

```
rad2degree(x)
```

Arguments

`x` an angle in radians

Value

the angle with the positive x-axis in degrees.

Author(s)

Jan Graffelman (jan.graffelman@upc.edu)

Examples

```
x <- pi/2
a <- rad2degree(x)
cat("angle is",a,"degrees\n")
```

rda *Redundancy analysis*

Description

rda performs redundancy analysis and stores extensive output in a list object.

Usage

```
rda(X, Y, scaling = 1)
```

Arguments

X	a matrix of x variables
Y	a matrix of y variables
scaling	scaling used for x and y variables. 0: x and y only centered. 1: x and y standardized

Details

Results are computed by doing a principal component analysis of the fitted values of the regression of y on x.

Plotting the first two columns of Gxs and Gyp, or of Gxp and Gys provides a biplots of the matrix of regression coefficients.

Plotting the first two columns of Fs and Gp or of Fp and Gs provides a biplot of the matrix of fitted values.

Value

Returns a list with the following results

Yh	fitted values of the regression of y on x
B	regression coefficients of the regression of y on x
decom	variance decomposition/goodness of fit of the fitted values AND of the regression coefficients
Fs	biplot markers of the rows of Yh (standard coordinates)
Fp	biplot markers of the rows of Yh (principal coordinates)
Gys	biplot markers for the y variables (standard coordinates)
Gyp	biplot markers for the y variables (principal coordinates)
Gxs	biplot markers for the x variables (standard coordinates)
Gxp	biplot markers for the x variables (principal coordinates)

Author(s)

Jan Graffelman (jan.graffelman@upc.edu)

References

Van den Wollenberg, A.L. (1977) Redundancy Analysis, an alternative for canonical correlation analysis. *Psychometrika* 42(2): pp. 207-219.

Ter Braak, C. J. F. and Looman, C. W. N. (1994) Biplots in Reduced-Rank Regression. *Biometrical Journal* 36(8): pp. 983-1003.

See Also

[princomp](#), [canocor](#), [biplot](#)

Examples

```
X <- matrix(rnorm(75),ncol=3)
Y <- matrix(rnorm(75),ncol=3)
rda.results <- rda(X,Y)
```

shiftvector

Compute a shift vector for a calibrated axis.

Description

shiftvector computes two shift vectors perpendicular to the supplied biplot or scatterplot axis g. The vector norm is computed from the two most extreme data points.

Usage

```
shiftvector(g, X, x = c(1, 0), verbose = FALSE)
```

Arguments

g	a biplot or scatterplot axis
X	a n by 2 matrix of scatterplot or biplot coordinates
x	reference axis, (1,0) by default
verbose	print information or not

Details

shiftvector locates the tow most extreme datapoints in the direction perpendicular to axis g.

Value

dr	the right (w.r.t. the direction of g) shift vector
dl	the left (w.r.t. the direction of g) shift vector

Author(s)

Jan Graffelman (jan.graffelman@upc.edu)

References

Graffelman, J. and van Eeuwijk, F.A. (2005) Calibration of multivariate scatter plots for exploratory analysis of relations within and between sets of variables in genomic research *Biometrical Journal*, 47(6) pp. 863-879.

Graffelman, J. (2006) A guide to biplot calibration.

See Also

[calibrate](#)

Examples

```
X <- matrix(rnorm(100), ncol=2)
Xs <- scale(X)

g <- c(1,1)

plot(Xs[,1], Xs[,2], asp=1, pch=19)
textxy(Xs[,1], Xs[,2], 1:nrow(X))

arrows(0,0,g[1],g[2])
text(g[1],g[2], "g", pos=1)

out <- shiftvector(g,X,verbose=TRUE)
dr <- out$dr
dl <- out$dl

arrows(0,0,dl[1],dl[2])
text(dl[1],dl[2], "dl", pos=1)

arrows(0,0,dr[1],dr[2])
text(dr[1],dr[2], "dr", pos=1)
```

spaindist

Road distances between Spanish cities

Description

Road distances in kilometers between 47 Spanish cities

Usage

```
data(spaindist)
```

Format

A data frame containing 47 observations.

References

Graffelman, J. (2019) Goodness-of-fit filtering in classical metric multidimensional scaling with large datasets. <doi: 10.1101/708339>

storks

Frequencies of nesting storks in Denmark

Description

Danish data from 1953-1977 giving the frequency of nesting storks, the human birth rate and the per capita electricity consumption.

Usage

data(storks)

Format

A data frame containing 25 observations.

Source

Gabriel and Odoroff, Table 1.

References

Gabriel, K. R. and Odoroff, C. L. (1990) Biplots in biomedical research. *Statistics in Medicine* 9(5): pp. 469-485.

textxy

Nice placement of labels in a plot

Description

Function textxy calls function text in order to add text to points in a graph. textxy chooses a different position for the text depending on the quadrant. This tends to produce better readable plots, with labels fanning away from the origin.

Usage

textxy(X, Y, labs, m = c(0, 0), cex = 0.5, offset = 0.8, ...)

Arguments

X	x coordinates of a set of points
Y	y coordinates of a set of points
labs	labels to be placed next to the points
m	coordinates of the origin of the plot (default (0,0))
cex	character expansion factor
offset	controls the distance between the label and the point. A value of 0 will plot labels on top of the point. Larger values give larger separation between point and label. The default value is 0.8
...	additional arguments for function text.

Value

NULL

Author(s)

Jan Graffelman (jan.graffelman@upc.edu)

References

Graffelman, J. (2006) A guide to biplot calibration.

See Also

[text](#)

Examples

```
x <- rnorm(50)
y <- rnorm(50)
plot(x,y,asp=1)
textxy(x,y,1:50,m=c(mean(x),mean(y)))
```

Index

- *Topic **aplot**
 - circle, 8
 - dlines, 9
 - textxy, 18
- *Topic **arith**
 - rad2degree, 14
- *Topic **datasets**
 - calves, 6
 - goblets, 10
 - heads, 10
 - linnerud, 11
 - spaindist, 17
 - storks, 18
- *Topic **misc**
 - textxy, 18
- *Topic **multivariate**
 - bplot, 2
 - calibrate, 3
 - canocor, 6
 - ones, 11
 - origin, 12
 - PrinCoor, 13
 - rda, 15
 - shiftvector, 16
- biplot, 5, 16
- bplot, 2
- calibrate, 3, 17
- calves, 6
- cancor, 7
- canocor, 6, 16
- circle, 8
- dlines, 9
- goblets, 10
- heads, 10
- lines, 9, 12
- linnerud, 11
- matrix, 12
- ones, 11
- origin, 12
- princomp, 14, 16
- PrinCoor, 13
- rad2degree, 14
- rda, 15
- shiftvector, 16
- spaindist, 17
- storks, 18
- text, 19
- textxy, 18