

Package ‘Correplot’

February 19, 2015

Type Package

Title A collection of functions for graphing correlation matrices

Version 1.0-2

Date 2013-10-12

Author Jan Graffelman

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

Depends R (>= 1.8.0), calibrate

Imports xtable

Description Correplot contains diverse routines for the construction of different plots for representing correlation matrices.

License GPL (>= 2)

URL <http://www.r-project.org>, <http://www-eio.upc.es/~jan>

NeedsCompilation no

Repository CRAN

Date/Publication 2013-10-13 19:40:50

R topics documented:

aircraftR	2
angleToR	3
artificialR	4
athletesR	4
berkeleyR	5
cathedralsR	5
correlogram	6
countriesR	7
fit_angles	7
physiologyR	8
gobletsR	9
linangplot	9
lincos	10

pco	11
PearsonLee	12
pfa	13
proteinR	14
proteinsR	15
recordsR	15
storksR	16
students	16
studentsR	17
Index	18

aircraftR	<i>Correlations between characteristics of aircraft</i>
-----------	---------------------------------------------------------

Description

Correlations between SPR (specific power), RGF (flight range factor), PLF (payload) and SLF (sustained load factor) for 21 types of aircraft.

Usage

```
data(aircraftR)
```

Format

a matrix containing the correlations

Source

Gower and Hand, Table 2.1

References

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

angleToR *Convert angles to correlations.*

Description

Function `angleToR` converts a vector of angles (in radians) to an estimate of the correlation matrix, given an interpretation function.

Usage

```
angleToR(x, ifun = "cos")
```

Arguments

`x` a vector of angles (in radians)
`ifun` the interpretation function ("cos" or "lincos")

Value

A correlation matrix

Author(s)

Jan Graffelman (jan.graffelman@upc.edu)

References

Graffelman, J. (2012) Linear-angle correlation plots: new graphs for revealing correlation structure. *Journal of Computational and Graphical Statistics*. 22(1): 92-106.

See Also

[cos](#), [lincos](#)

Examples

```
angles <- c(0, pi/3)
R <- angleToR(angles)
print(R)
```

artificialR

Correlations for 10 generated variables

Description

A 10 by 10 artificial correlation matrix

Usage

```
data(artificialR)
```

Format

A matrix of correlations

Source

Trosset (2005), Table 1.

References

Trosset, M.W. (2005) Visualizing correlation. *Journal of Computational and Graphical Statistics*, 14(1), pp. 1–19.

athletesR

Correlation matrix of characteristics of Australian athletes

Description

Correlation matrix of 12 characteristics of Austration athletes (Sex, Height, Weight, Lean Body Mass, RCC, WCC, Hc, Hg, Ferr, BMI, SSF, Bfat)

Usage

```
data(athletesR)
```

Format

A matrix of correlations

Source

Weisberg (2005), file ais.txt

References

Weisberg, S. (2005) *Applied Linear Regression*. Third edition, John Wiley & Sons, New Jersey.

`berkeleyR`*Correlation matrix for boys of the Berkeley Guidance Study*

Description

Correlation matrix for sex, height and weight at age 2, 9 and 18 and somatotype

Usage

```
data(berkeleyR)
```

Format

A matrix of correlations

Source

Weisberg (2005), file BGSBoys.txt

References

Weisberg, S. (2005) *Applied Linear Regression*. Third edition, John Wiley & Sons, New Jersey.

`cathedralsR`*Correlation matrix for height and length*

Description

Correlation between nave height and total length

Usage

```
data(cathedralsR)
```

Format

A matrix of correlations

Source

Weisberg (2005), file cathedral.txt

References

Weisberg, S. (2005) *Applied Linear Regression*. Third edition, John Wiley & Sons, New Jersey.

`correlogram`*Plot a correlogram*

Description

`correlogram` plots a correlogram for a correlation matrix.

Usage

```
correlogram(R, labs=colnames(R), ifun="cos", cex=1, main="", ntrials=10,  
            xlim=c(-1.2, 1.2), ylim=c(-1.2, 1.2), ...)
```

Arguments

<code>R</code>	a correlation matrix.
<code>labs</code>	a vector of labels for the variables.
<code>ifun</code>	the interpretation function ("cos" or "lincos")
<code>cex</code>	character expansion factor for the variable labels
<code>main</code>	a title for the correlogram
<code>ntrials</code>	number of starting points for the optimization routine
<code>xlim</code>	limits for the x axis (e.g. <code>c(-1.2, 1.2)</code>)
<code>ylim</code>	limits for the y axis (e.g. <code>c(-1.2, 1.2)</code>)
<code>...</code>	additional arguments for the <code>plot</code> function.

Details

`correlogram` makes a correlogram on the basis of a set of angles. All angles are given w.r.t the positive x-axis. Variables are represented by unit vectors emanating from the origin.

Value

NULL

Author(s)

anonymous

References

Trosset, M.W. (2005) Visualizing correlation. *Journal of Computational and Graphical Statistics* 14(1), pp. 1–19

See Also

[fit_angles](#), [nlminb](#)

Examples

```
X <- matrix(rnorm(90),ncol=3)
R <- cor(X)
correlogram(R)
```

countriesR

Correlations between educational and demographic variables

Description

Correlations between infant mortality, educational and demographic variables (infd, phys, dens, agds, lit, hied, gnp)

Usage

```
data(countriesR)
```

Format

A matrix of correlations

Source

Chatterjee and Hadi (1988)

References

Chatterjee, S. and Hadi, A.S. (1988), *Sensitivity Analysis in Regression*. Wiley, New York.

fit_angles

Fit angles to a correlation matrix

Description

fit_angles finds a set of optimal angles for representing a particular correlation matrix by angles between vectors

Usage

```
fit_angles(R, ifun = "cos", ntrials = 10, verbose = FALSE)
```

Arguments

R	a correlation matrix.
ifun	an angle interpretation function (cosine, by default).
ntrials	number of trials for optimization routine nlmnb
verbose	be silent (FALSE), or produce more output (TRUE)

Value

a vector of angles (in radians)

Author(s)

anonymous

References

Trosset, M.W. (2005) Visualizing correlation. *Journal of Computational and Graphical Statistics* 14(1), pp. 1–19

See Also

[nlminb](#)

Examples

```
X <- matrix(rnorm(90), ncol=3)
R <- cor(X)
angles <- fit_angles(R)
print(angles)
```

fysiologyR

Correlations between thirteen physiological variables

Description

Correlations of 13 physiological variables (sys, dia, p.p., pul, cort, u.v., tot/100, adr/100, nor/100, adr/tot, tot/hr, adr/hr, nor/hr) obtained from 48 medical students

Usage

```
data(fysiologyR)
```

Format

A matrix of correlations

Source

Hills (1969), Table 1.

References

Hills, M (1969) On looking at large correlation matrices *Biometrika* 56(2): pp. 249.

`gobletsR`*Correlations between size measurements of archeological goblets*

Description

Correlations between 6 size measurements of archeological goblets

Usage

```
data(gobletsR)
```

Format

A matrix of correlations

Source

Manly (1989)

References

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

`linangplot`*Linang plot*

Description

`linangplot` produces a plot of two variables, such that the correlation between the two variables is linear in the angle.

Usage

```
linangplot(x, y, tmx = NULL, tmy = NULL, ...)
```

Arguments

<code>x</code>	x variable
<code>y</code>	y variable
<code>tmx</code>	vector of tickmarks for the x variable
<code>tmy</code>	vector of tickmarks for the y variable
<code>...</code>	additional arguments for the plot routine

Value

Xt	coordinates of the points
B	axes for the plot
r	correlation coefficient
angledegrees	angle between axes in degrees
angleradians	angle between axes in radians
r	correlation coefficient

Author(s)

Jan Graffelman (jan.graffelman@upc.edu)

See Also

[plotcorrelogram](#)

Examples

```
x <- runif(10)
y <- rnorm(10)
linangplot(x,y)
```

lincos

Linearized cosine function

Description

Function `lincos` linearizes the cosine function over the interval $[0, 2\pi]$. The function returns $-2/\pi * x + 1$ over $[0, \pi]$ and $2/\pi * x - 3$ over $[\pi, 2\pi]$

Usage

```
lincos(x)
```

Arguments

x angle in radians

Value

a real number in $[-1, 1]$.

Author(s)

Jan Graffelman (jan.graffelman@upc.edu)

References

Graffelman, J. (2012) Linear-angle correlation plots: new graphs for revealing correlation structure. *Journal of Computational and Graphical Statistics*. 22(1): 92-106.

See Also

[cos](#)

Examples

```
angle <- pi
y <- lincos(angle)
print(y)
```

pco

Principal Coordinate Analysis

Description

pco is a program for Principal Coordinate Analysis.

Usage

```
pco(Dis)
```

Arguments

Dis A distance or dissimilarity matrix

Details

The program pco does a principal coordinates analysis of a dissimilarity (or distance) matrix (Dij) where the diagonal elements, Dii, are zero.

Note that when we dispose of a similarity matrix rather than a distance matrix, a transformation is needed before calling `coorprincipal`. For instance, if Sij is a similarity matrix, Dij might be obtained as $D_{ij} = 1 - S_{ij}/\text{diag}(S_{ij})$

Goodness of fit calculations need to be revised such as to deal (in different ways) with negative eigenvalues.

Value

PC	the principal coordinates
D1	all eigenvalues of the solution
Dk	the positive eigenvalues of the solution
B	double centred matrix for the eigenvalue decomposition
decom	the goodness of fit table

Author(s)

Jan Graffelman (jan.graffelman@upc.edu)

See Also

[cmdscale](#)

Examples

```
citynames <- c("Aberystwyth", "Brighton", "Carlisle", "Dover", "Exeter", "Glasgow", "Hull",
              "Inverness", "Leeds", "London", "Newcastle", "Norwich")
A <-matrix(c(
0,244,218,284,197,312,215,469,166,212,253,270,
244,0,350,77,167,444,221,583,242,53,325,168,
218,350,0,369,347,94,150,251,116,298,57,284,
284,77,369,0,242,463,236,598,257,72,340,164,
197,167,347,242,0,441,279,598,269,170,359,277,
312,444,94,463,441,0,245,169,210,392,143,378,
215,221,150,236,279,245,0,380,55,168,117,143,
469,583,251,598,598,169,380,0,349,531,264,514,
166,242,116,257,269,210,55,349,0,190,91,173,
212,53,298,72,170,392,168,531,190,0,273,111,
253,325,57,340,359,143,117,264,91,273,0,256,
270,168,284,164,277,378,143,514,173,111,256,0),ncol=12)
rownames(A) <- citynames
colnames(A) <- citynames
out <- pco(A)
plot(out$PC[,2],-out$PC[,1],pch=19,asp=1)
textxy(out$PC[,2],-out$PC[,1],rownames(A))
```

PearsonLee

Heights of mothers and daughters

Description

Heights of 1375 mothers and daughters (in cm) in the UK in 1893-1898.

Usage

```
data(PearsonLee)
```

Format

dataframe with Mheight and Dheight

Source

Weisberg, Chapter 1

References

Weisberg, S. (2005) *Applied Linear Regression*, John Wiley & Sons, New Jersey

pfa *Principal factor analysis*

Description

Program pfa performs (iterative) principal factor analysis, which is based on the computation of eigenvalues of the reduced correlation matrix.

Usage

```
pfa(X, option = "data", m = 2, initial.communality = "R2", crit = 0.001, verbose = FALSE)
```

Arguments

X	A data matrix or correlation matrix
option	Specifies the type of matrix supplied by argument X. Values for option are data, cor or cov. data is the default.
m	The number of factors to extract (2 by default)
initial.communality	Method for computing initial communalities. Possibilities are R2 or maxcor.
crit	The criterion for convergence. The default is 0.001. A smaller value will require more iterations before convergence is reached.
verbose	When set to TRUE, additional numerical output is shown.

Value

Res	Matrix of residuals
Psi	Diagonal matrix with specific variances
La	Matrix of loadings
Shat	Estimated correlation matrix
Fs	Factor scores

Author(s)

Jan Graffelman (jan.graffelman@upc.edu)

References

Mardia, K.V., Kent, J.T. and Bibby, J.M. (1979) *Multivariate analysis*.

Rencher, A.C. (1995) *Methods of multivariate analysis*.

Satorra, A. and Neudecker, H. (1998) Least-Squares Approximation of off-Diagonal Elements of a Variance Matrix in the Context of Factor Analysis. *Econometric Theory* 14(1) pp. 156–157.

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

```
X <- matrix(rnorm(100),ncol=2)
out.pfa <- pfa(X)
# based on a correlation matrix
R <- cor(X)
out.pfa <- pfa(R,option="cor")
```

proteinR

Correlations between sources of protein

Description

Correlations between sources of protein for a number of countries (Red meat, White meat, Eggs, Milk, Fish, Cereals, Starchy food, Nuts, Fruits and vegetables).

Usage

```
data(proteinR)
```

Format

A matrix of correlations

Source

Manly (1989)

References

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

proteinsR	<i>Correlations between sources of protein</i>
-----------	------------------------------------------------

Description

Correlations between sources of protein for a number of countries (Red meat, White meat, Eggs, Milk, Fish, Cereals, Starchy food, Nuts, Fruits and vegetables).

Usage

```
data(proteinR)
```

Format

A matrix of correlations

Source

Manly (1989)

References

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

recordsR	<i>Correlations between national track records for men</i>
----------	------------------------------------------------------------

Description

Correlations between national track records for men (100m,200m,400m,800m,1500m,5000m,10.000m and Marathon)

Usage

```
data(recordsR)
```

Format

A matrix of correlations

Source

Johnson and Wichern, Table 8.6

References

Johnson, R.A. and Wichern, D.W. (2002) *Applied Multivariate Statistical Analysis*. Fifth edition. New Jersey: Prentice Hall.

storksR

Correlations between three variables

Description

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

Usage

```
data(storksR)
```

Format

A matrix of correlations

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.

students

Marks for 5 student exams

Description

Matrix of marks for five exams, two with closed books and three with open books (Mechanics (C), Vectors (C), Algebra (O), Analysis (O) and Statistics (O)).

Usage

```
data(students)
```

Format

A data matrix

Source

Mardia et al., Table 1.2.1

References

Mardia, K.V., Kent, J.T. and Bibby, J.M. (1979) *Multivariate Analysis*, Academic Press London.

`studentsR`*Correlations between marks for 5 exams*

Description

Correlation matrix of marks for five exams, two with closed books and three with open books (Mechanics (C), Vectors (C), Algebra (O), Analysis (O) and Statistics (O)).

Usage

```
data(studentsR)
```

Format

A matrix of correlations

Source

Mardia et al., Table 1.2.1

References

Mardia, K.V., Kent, J.T. and Bibby, J.M. (1979) *Multivariate Analysis*, Academic Press London.

Index

*Topic **datasets**

- aircraftR, [2](#)
- artificialR, [4](#)
- athletesR, [4](#)
- berkeleyR, [5](#)
- cathedralsR, [5](#)
- countriesR, [7](#)
- fysiologyR, [8](#)
- gobletsR, [9](#)
- PearsonLee, [12](#)
- proteinR, [14](#)
- proteinsR, [15](#)
- recordsR, [15](#)
- storksR, [16](#)
- students, [16](#)
- studentsR, [17](#)

*Topic **hplot**

- correlogram, [6](#)
- linangplot, [9](#)

*Topic **misc**

- angleToR, [3](#)
- lincos, [10](#)

*Topic **multivariate**

- pco, [11](#)
- pfa, [13](#)

*Topic **optimize**

- fit_angles, [7](#)

- aircraftR, [2](#)
- angleToR, [3](#)
- artificialR, [4](#)
- athletesR, [4](#)

- berkeleyR, [5](#)

- cathedralsR, [5](#)
- cmdscale, [12](#)
- correlogram, [6](#), [10](#)
- cos, [3](#), [11](#)
- countriesR, [7](#)

- fit_angles, [6](#), [7](#)

- fysiologyR, [8](#)

- gobletsR, [9](#)

- linangplot, [9](#)

- lincos, [3](#), [10](#)

- nlminb, [6](#), [8](#)

- pco, [11](#)

- PearsonLee, [12](#)

- pfa, [13](#)

- plot, [10](#)

- princomp, [14](#)

- proteinR, [14](#)

- proteinsR, [15](#)

- recordsR, [15](#)

- storksR, [16](#)

- students, [16](#)

- studentsR, [17](#)