Package 'MFAg'

May 22, 2020

Type Package

Title Multiple Factor Analysis (MFA)

Version 1.7

Date 2020-05-21

Author Paulo Cesar Ossani <ossanipc@hotmail.com>
Marcelo Angelo Cirillo <macufla@dex.ufla.br>

Maintainer Paulo Cesar Ossani <ossanipc@hotmail.com>

Description Performs Multiple Factor Analysis method for quantitative, categorical, frequency and mixed data, in addition to generating a lot of graphics, also has other useful functions.

License GPL (>= 2)

NeedsCompilation no

Repository CRAN

Date/Publication 2020-05-22 12:40:02 UTC

R topics documented:

DataMix *Mixed data set.*

Description

Simulated set of mixed data on consumption of coffee.

Usage

```
data(DataMix)
```

Format

Data set with 10 rows and 7 columns. Being 10 observations described by 7 variables: Cooperatives/Tasters, Average grades given to analyzed coffees, Years of work as a taster, Taster with technical training, Taster exclusively dedicated, Average frequency of the coffees Classified as special, Average frequency of the coffees as commercial.

Author(s)

Paulo Cesar Ossani

Marcelo Angelo Cirillo

Examples

```
data(DataMix)
DataMix
```

DataQuali *Qualitative data set*

Description

Set simulated of qualitative data on consumption of coffee.

Usage

```
data(DataQuali)
```

Format

Data set simulated with 12 rows and 6 columns. Being 12 observations described by 6 variables: Sex, Age, Smoker, Marital status, Sportsman, Study.

Author(s)

Paulo Cesar Ossani

Marcelo Angelo Cirillo

Examples

```
data(DataQuali)
DataQuali
```

`DataQuan` *Quantitative data set*

Description

Set simulated of quantitative data on grades given to some sensory characteristics of coffees.

Usage

```
data(DataQuan)
```

Format

Data set with 6 rows and 11 columns. Being 6 observations described by 11 variables: Coffee, Chocolate, Caramelised, Ripe, Sweet, Delicate, Nutty, Caramelised, Chocolate, Spicy, Caramelised.

Author(s)

Paulo Cesar Ossani

Marcelo Angelo Cirillo

Examples

```
data(DataQuan)
DataQuan
```

GSVD *Generalized Singular Value Decomposition (GSVD).*

Description

Given the matrix A of order nxm, the generalized singular value decomposition (GSVD) involves the use of two sets of positive square matrices of order nxn and mxm respectively. These two matrices express constraints imposed, respectively, on the lines and columns of A.

Usage

```
GSVD(data, plin = NULL, pcol = NULL)
```

Arguments

data	Matrix used for decomposition.
plin	Weight for rows.
pcol	Weight for columns

Details

If plin or pcol is not used, it will be calculated as the usual singular value decomposition.

Value

d	Eigenvalues, that is, line vector with singular values of the decomposition.
u	Eigenvectors referring rows.
v	Eigenvectors referring columns.

Author(s)

Paulo Cesar Ossani

Marcelo Angelo Cirillo

References

ABDI, H. Singular Value Decomposition (SVD) and Generalized Singular Value Decomposition (GSVD). In: SALKIND, N. J. (Ed.). *Encyclopedia of measurement and statistics.* Thousand Oaks: Sage, 2007. p. 907-912.

Examples

```
M = matrix(c(1,2,3,4,5,6,7,8,9,10,11,12), nrow = 4, ncol = 3)

svd(M)  # Usual Singular Value Decomposition

GSVD(M) # GSVD with the same previous results

# GSVD with weights for rows and columns
GSVD(M, plin = c(0.1,0.5,2,1.5), pcol = c(1.3,2,0.8))
```

IM *Indicator matrix.*

Description

In the indicator matrix the elements are arranged in the form of *dummy* variables, in other words, 1 for a category chosen as a response variable and 0 for the other categories of the same variable.

Usage

```
IM(data, names = TRUE)
```

Arguments

data	Categorical data.
names	Include the names of the variables in the levels of the Indicator Matrix (default = TRUE).

Value

mtxIndc	Returns converted data in the indicator matrix.

Author(s)

Paulo Cesar Ossani

Marcelo Angelo Cirillo

References

RENCHER, A. C. *Methods of multivariate analysis.* 2th. ed. New York: J.Wiley, 2002. 708 p.

Examples

```
data <- matrix(c("S","S","N","N",1,2,3,4,"N","S","T","N"), nrow = 4, ncol = 3)

IM(data, names = FALSE)

data(DataQuali) # qualitative data set

IM(DataQuali, names = TRUE)
```

LocLab *Function for better position of the labels in the graphs.*

Description

Function for better position of the labels in the graphs.

Usage

```
LocLab(x, y = NULL, labels = seq(along = x), cex = 1,
       method = c("SANN", "GA"), allowSmallOverlap = FALSE,
       trace = FALSE, shadotext = FALSE,
       doPlot = TRUE, ...)
```

Arguments

x	Coordinate x
y	Coordinate y
labels	The labels
cex	cex
method	Not used
allowSmallOverlap	Boolean
trace	Boolean
shadotext	Boolean
doPlot	Boolean
...	Other arguments passed to or from other methods

Value

See the text of the function.

MFA *Multiple Factor Analysis (MFA).*

Description

Perform Multiple Factor Analysis (MFA) on groups of variables. The groups of variables can be quantitative, qualitative, frequency (MFACT) data, or mixed data.

Usage

```
MFA(data, groups, typegroups = rep("n",length(groups)), namegroups = NULL)
```

Arguments

data	Data to be analyzed.
groups	Number of columns for each group in order following the order of data in 'data'.
typegroups	Type of group: "n" for numerical data (default), "c" for categorical data, "f" for frequency data.
namegroups	Names for each group.

Value

vtrG	Vector with the sizes of each group.
vtrNG	Vector with the names of each group.
vtrplin	Vector with the values used to balance the lines of the Z matrix.
vtrpcol	Vector with the values used to balance the columns of the Z matrix.
mtxZ	Matrix concatenated and balanced.
mtxA	Matrix of the eigenvalues (variances) with the proportions and proportions accumulated.
mtxU	Matrix U of the singular decomposition of the matrix Z.
mtxV	Matrix V of the singular decomposition of the matrix Z.
mtxF	Matrix global factor scores where the lines are the observations and the columns the components.
mtxEFG	Matrix of the factor scores by group.
mtxCCP	Matrix of the correlation of the principal components with original variables.
mtxEV	Matrix of the partial inertias / scores of the variables

Author(s)

Paulo Cesar Ossani

Marcelo Angelo Cirillo

References

ABDESSEMED, L.; ESCOFIER, B. Analyse factorielle multiple de tableaux de frequencies: comparaison avec l'analyse canonique des correspondences. *Journal de la Societe de Statistique de Paris*, Paris, v. 137, n. 2, p. 3-18, 1996..

ABDI, H. Singular Value Decomposition (SVD) and Generalized Singular Value Decomposition (GSVD). In: SALKIND, N. J. (Ed.). *Encyclopedia of measurement and statistics.* Thousand Oaks: Sage, 2007. p. 907-912.

ABDI, H.; VALENTIN, D. Multiple factor analysis (MFA). In: SALKIND, N. J. (Ed.). *Encyclopedia of measurement and statistics.* Thousand Oaks: Sage, 2007. p. 657-663.

ABDI, H.; WILLIAMS, L. Principal component analysis. *WIREs Computational Statatistics*, New York, v. 2, n. 4, p. 433-459, July/Aug. 2010.

ABDI, H.; WILLIAMS, L.; VALENTIN, D. Multiple factor analysis: principal component analysis for multitable and multiblock data sets. *WIREs Computational Statatistics*, New York, v. 5, n. 2, p. 149-179, Feb. 2013.

BECUE-BERTAUT, M.; PAGES, J. A principal axes method for comparing contingency tables: MFACT. *Computational Statistics & data Analysis*, New York, v. 45, n. 3, p. 481-503, Feb. 2004

BECUE-BERTAUT, M.; PAGES, J. Multiple factor analysis and clustering of a mixture of quantitative, categorical and frequency data. *Computational Statistics & data Analysis*, New York, v. 52, n. 6, p. 3255-3268, Feb. 2008.

BENZECRI, J. Analyse de l'inertie intraclasse par l'analyse d'un tableau de contingence: intra-classinertia analysis through the analysis of a contingency table. *Les Cahiers de l'Analyse des Donnees*, Paris, v. 8, n. 3, p. 351-358, 1983.

ESCOFIER, B. Analyse factorielle en reference a un modele: application a l'analyse d'un tableau d'echanges. *Revue de Statistique Appliquee*, Paris, v. 32, n. 4, p. 25-36, 1984.

ESCOFIER, B.; DROUET, D. Analyse des differences entre plusieurs tableaux de frequence. *Les Cahiers de l'Analyse des Donnees*, Paris, v. 8, n. 4, p. 491-499, 1983.

ESCOFIER, B.; PAGES, J. *Analyse factorielles simples et multiples.* Paris: Dunod, 1990. 267 p.

ESCOFIER, B.; PAGES, J. *Analyses factorielles simples et multiples:* objectifs, methodes et interpretation. 4th ed. Paris: Dunod, 2008. 318 p.

ESCOFIER, B.; PAGES, J. *Comparaison de groupes de variables definies sur le meme ensemble d'individus:* un exemple d'applications. Le Chesnay: Institut National de Recherche en Informatique et en Automatique, 1982. 121 p.

ESCOFIER, B.; PAGES, J. Multiple factor analysis (AFUMULT package). *Computational Statistics & data Analysis*, New York, v. 18, n. 1, p. 121-140, Aug. 1994

GREENACRE, M.; BLASIUS, J. *Multiple correspondence analysis and related methods.* New York: Taylor and Francis, 2006. 607 p.

OSSANI, P. C.; CIRILLO, M. A.; BOREM, F. M.; RIBEIRO, D. E.; CORTEZ, R. M.. Quality of specialty coffees: a sensory evaluation by consumers using the MFACT technique. *Revista Ciencia Agronomica (UFC. Online)*, v. 48, p. 92-100, 2017.

PAGES, J. Analyse factorielle multiple appliquee aux variables qualitatives et aux donnees mixtes. *Revue de Statistique Appliquee*, Paris, v. 50, n. 4, p. 5-37, 2002.

PAGES, J.. Multiple factor analysis: main features and application to sensory data. *Revista Colombiana de Estadistica*, Bogota, v. 27, n. 1, p. 1-26, 2004.

See Also

Plot.MFA

Examples

```
data(DataMix) # mixed dataset

data <- DataMix[,2:ncol(DataMix)]

rownames(data) <- DataMix[1:nrow(DataMix),1]
```

```
GroupNames = c("Grade Cafes/Work", "Formation/Dedication", "Coffees")

MF <- MFA(data = data, c(2,2,2), typegroups = c("n","c","f"), GroupNames) # performs MFA

print("Principal Component Variances:"); round(MF$mtxA,2)

print("Matrix of the Partial Inertia / Score of the Variables:"); round(MF$mtxEV,2)
```

MFAg *Multiple Factor Analysis (MFA)*

Description

Performs multiple factor analysis method for quantitative, categorical, frequency and mixed data.

Details

Package:	MFAg
Type:	Package
Version:	1.7
Date:	2020-05-21
License:	GPL (>=2)
LazyLoad:	yes

Author(s)

Paulo Cesar Ossani,

Marcelo Angelo Cirillo

Maintainer: Paulo Cesar Ossani <ossanipc@hotmail.com>

References

ABDESSEMED, L. and ESCOFIER, B.; Analyse factorielle multiple de tableaux de frequencies: comparaison avec l'analyse canonique des correspondences. *Journal de la Societe de Statistique de Paris*, Paris, v. 137, n. 2, p. 3-18, 1996.

ABDI, H. Singular Value Decomposition (SVD) and Generalized Singular Value Decomposition (GSVD). In: SALKIND, N. J. (Ed.). *Encyclopedia of measurement and statistics.* Thousand Oaks: Sage, 2007. p. 907-912.

ABDI, H.; VALENTIN, D. Multiple factor analysis (MFA). In: SALKIND, N. J. (Ed.). *Encyclopedia of measurement and statistics.* Thousand Oaks: Sage, 2007. p. 657-663.

ABDI, H.; WILLIAMS, L. Principal component analysis. *WIREs Computational Statatistics*, New York, v. 2, n. 4, p. 433-459, July/Aug. 2010.

ABDI, H.; WILLIAMS, L.; VALENTIN, D. Multiple factor analysis: principal component analysis for multitable and multiblock data sets. *WIREs Computational Statatistics*, New York, v. 5, n. 2, p. 149-179, Feb. 2013.

BECUE-BERTAUT, M.; PAGES, J. A principal axes method for comparing contingency tables: MFACT. *Computational Statistics & Data Analysis*, New York, v. 45, n. 3, p. 481-503, Feb. 2004

BECUE-BERTAUT, M.; PAGES, J. Multiple factor analysis and clustering of a mixture of quantitative, categorical and frequency data. *Computational Statistics & Data Analysis*, New York, v. 52, n. 6, p. 3255-3268, Feb. 2008.

BENZECRI, J. Analyse de l'inertie intraclasse par l'analyse d'un tableau de contingence: intra-classinertia analysis through the analysis of a contingency table. *Les Cahiers de l'Analyse des Donnees*, Paris, v. 8, n. 3, p. 351-358, 1983.

ESCOFIER, B. Analyse factorielle en reference a un modele: application a l'analyse d'un tableau d'echanges. *Revue de Statistique Appliquee*, Paris, v. 32, n. 4, p. 25-36, 1984.

ESCOFIER, B.; DROUET, D. Analyse des differences entre plusieurs tableaux de frequence. *Les Cahiers de l'Analyse des Donnees*, Paris, v. 8, n. 4, p. 491-499, 1983.

ESCOFIER, B.; PAGES, J. *Analyse factorielles simples et multiples.* Paris: Dunod, 1990. 267 p.

ESCOFIER, B.; PAGES, J. *Analyses factorielles simples et multiples:* objectifs, methodes et interpretation. 4th ed. Paris: Dunod, 2008. 318 p.

ESCOFIER, B.; PAGES, J. *Comparaison de groupes de variables definies sur le meme ensemble d'individus:* un exemple d'applications. Le Chesnay: Institut National de Recherche en Informatique et en Automatique, 1982. 121 p.

ESCOFIER, B.; PAGES, J. Multiple factor analysis (AFUMULT package). *Computational Statistics & Data Analysis*, New York, v. 18, n. 1, p. 121-140, Aug. 1994

FERREIRA, D. F. *Estatistica multivariada.* 2. ed. rev. e ampl. Lavras: UFLA, 2011. 675 p.

GREENACRE, M.; BLASIUS, J. *Multiple correspondence analysis and related methods.* New York: Taylor and Francis, 2006. 607 p.

HOTELLING, H. Analysis of a complex of statistical variables into principal components. *Journal of Educational Psychology*, Arlington, v. 24, p. 417-441, Sept. 1933.

JOHNSON, R. A.; WICHERN, D. W. *Applied multivariate statistical analysis.* 6th ed. New Jersey: Prentice Hall, 2007. 794 p.

MINGOTI, S. A. *Analise de dados atraves de metodos de estatistica multivariada:* uma abordagem aplicada. Belo Horizonte: UFMG, 2005. 297 p.

PAGES, J. Analyse factorielle multiple appliquee aux variables qualitatives et aux donnees mixtes. *Revue de Statistique Appliquee*, Paris, v. 50, n. 4, p. 5-37, 2002.

PAGES, J. Multiple factor analysis: main features and application to sensory data. *Revista Colombiana de Estadistica*, Bogota, v. 27, n. 1, p. 1-26, 2004.

OSSANI, P. C. *Qualidade de cafes especiais e nao especiais por meio da analise de multiplos fatores para tabelas de contingencias.* 2015. 107 p. Dissertacao (Mestrado em Estatistica e Experimentacao Agropecuaria) - Universidade Federal de Lavras, Lavras, 2015.

OSSANI, P. C.; CIRILLO, M. A.; Habilidades Sensoriais de Grupos Heterogeneos de Consumidores de cafes Especiais Discriminadas pelo Metodo MFACT. in: XIII ENCONTRO MINEIRO DE ESTATISTICA (MGEST), 13., 2014, Diamantina. *Anais*...Diamantina: UFVJM, 2014.

OSSANI, P. C. et al.; Multiplos fatores em analise de tabela de contingencia: Uma aplicacao na analise sensorial da qualidade de cafes especiais. in: 59 REUNIAO ANUAL DA REGIAO BRASILEIRA DA SOCIEDADE INTERNACIONAL DE BIOMETRIA (RBRAS), 59., 2014, Ouro Preto. *Anais*...Ouro Preto: UFOP, 2014.

RENCHER, A.C.; *Methods of Multivariate Analysis.* 2th. ed. New York: J.Wiley, 2002. 708 p.

`NormData` *Normalizes the data.*

Description

Function that normalizes the data globally, or by column.

Usage

```
NormData(data, type = 1)
```

Arguments

`data`	Data to be analyzed.
`type`	1 normalizes overall (default), 2 normalizes per column.

Value

`dataNorm`	Normalized data.

Author(s)

Paulo Cesar Ossani

Marcelo Angelo Cirillo

Examples

```
data(DataQuan) # set of quantitative data

data <- DataQuan[,2:8]

Resp = NormData(data, type = 1) # normalizes the data globally

Resp # Globally standardized data

sd(Resp)   # overall standard deviation

mean(Resp) # overall mean


Resp = NormData(data, type = 2) # normalizes the data per column
```

```
Resp # standardized data per column

apply(Resp, 2, sd) # standard deviation per column

colMeans(Resp)     # column averages
```

`Plot.MFA` *Graphics of the Multiple Factor Analysis (MFA).*

Description

Graphics of the Multiple Factor Analysis (MFA).

Usage

```
Plot.MFA(MFA, titles = NA, xlabel = NA, ylabel = NA,
         posleg = 2, boxleg = TRUE, size = 1.1, grid = TRUE,
         color = TRUE, groupscolor = NA, namarr = FALSE,
         linlab = NA, casc = TRUE)
```

Arguments

`MFA`	Data of the MFA function.
`titles`	Titles of the graphics, if not set, assumes the default text.
`xlabel`	Names the X axis, if not set, assumes the default text.
`ylabel`	Names the Y axis, if not set, assumes the default text.
`posleg`	1 for caption in the left upper corner, 2 for caption in the right upper corner (default), 3 for caption in the right lower corner, 4 for caption in the left lower corner.
`boxleg`	Puts frame in legend (default = TRUE).
`size`	Size of the points in the graphs.
`grid`	Put grid on graphs (default = TRUE).
`color`	Colored graphics (default = TRUE).
`groupscolor`	Vector with the colors of the groups.
`namarr`	Puts the points names in the cloud around the centroid in the graph corresponding to the global analysis of the Individuals and Variables (default = FALSE).
`linlab`	Vector with the labels for the observations, if not set, assumes the default text.
`casc`	Cascade effect in the presentation of the graphics (default = TRUE).

Value

Returns several graphs.

Author(s)

Paulo Cesar Ossani

Marcelo Angelo Cirillo

See Also

MFA

Examples

```
data(DataMix) # set of mixed data

Data <- DataMix[,2:ncol(DataMix)]

rownames(Data) <- DataMix[1:nrow(DataMix),1]

GroupNames = c("Grade Cafes/Work", "Formation/Dedication", "Coffees")

MF <- MFA(Data, c(2,2,2), typegroups = c("n","c","f"), GroupNames) # performs MFA

Tit = c("Scree-Plot","Observations","Observations/Variables","Inertia of the Variable Groups")

Plot.MFA(MF, titles = Tit, xlabel = NA, ylabel = NA,
         posleg = 2, boxleg = FALSE, color = TRUE,
         groupscolor = c("blue3","red","goldenrod3"),
         namarr = FALSE, linlab = NA,
         casc = FALSE) # plotting several graphs on the screen

Plot.MFA(MF, titles = Tit, xlabel = NA, ylabel = NA,
         posleg = 2, boxleg = FALSE, color = TRUE,
         namarr = FALSE, linlab = rep("A?",10),
         casc = FALSE) # plotting several graphs on the screen
```

Index