

Package ‘PCAmixdata’

October 17, 2016

Type Package

Title Multivariate Analysis of Mixed Data

Version 3.0

Date 2016-10-09

Author Marie Chavent and Vanessa Kuentz and Amaury Labenne and Benoit Lique
and Jerome Saracco.

Maintainer Marie Chavent <marie.chavent@u-bordeaux.fr>

Description Principal Component Analysis, orthogonal rotation and multiple
factor analysis for a mixture of quantitative and qualitative variables.

License GPL (>=2.0)

RoxygenNote 5.0.1

R topics documented:

decathlon	2
dogs	3
flower	3
gironde	4
Lg	5
Lg.pond	6
MFAmix	7
PCAmix	10
PCArrot	13
plot.MFAmix	15
plot.PCAmix	19
predict.MFAmix	21
predict.PCAmix	23
print.MFAmix	24
print.PCAmix	25
protein	25
recod	26
recodqual	27
recodquant	27

RV	28
RV.pond	29
sol.2dim	30
splitgroups	31
splitmix	31
summary.MFamix	32
summary.PCAmix	33
supvar	34
supvar.MFamix	34
supvar.PCAmix	35
svd.triplet	36
tab.disjonctif.NA	37
vnf	37
wine	38
Index	39

decathlon	<i>Performance in decathlon (data)</i>
-----------	--

Description

The data used here refer to athletes’ performance during two sporting events.

Usage

data(decathlon)

Format

A data frame with 41 rows and 13 columns: the first ten columns corresponds to the performance of the athletes for the 10 events of the decathlon. The columns 11 and 12 correspond respectively to the rank and the points obtained. The last column is a categorical variable corresponding to the sporting event (2004 Olympic Game or 2004 Decastar)

Source

The references below.

References

Departement of Applied Mathematics, Agrocampus Rennes.
Le, S., Josse, J. & Husson, F. (2008). FactoMineR: An R Package for Multivariate Analysis. Journal of Statistical Software. 25(1). pp. 1-18.

dogs

Breeds of Dogs data

Description

Data refering to 27 breeds of dogs.

Format

A data frame with 27 rows (the breeds of dogs) and 7 columns: their size, weight and speed with 3 categories (small, medium, large), their intelligence (low, medium, high), their affectivity and aggressiveness with 3 categories (low, high), their function (utility, compagny, hunting).

Source

Originated by A. Brefort (1982) and cited in Saporta G. (2011).

flower

Flower Characteristics

Description

8 characteristics for 18 popular flowers.

Usage

```
data(flower)
```

Format

A data frame with 18 observations on 8 variables:

[, "V1"]	factor	winters
[, "V2"]	factor	shadow
[, "V3"]	factor	tubers
[, "V4"]	factor	color
[, "V5"]	ordered	soil
[, "V6"]	ordered	preference
[, "V7"]	numeric	height
[, "V8"]	numeric	distance

V1 winters, is binary and indicates whether the plant may be left in the garden when it freezes.

V2 shadow, is binary and shows whether the plant needs to stand in the shadow.

V3 tubers, is asymmetric binary and distinguishes between plants with tubers and plants that grow in any other way.

- V4** color, is nominal and specifies the flower's color (1 = white, 2 = yellow, 3 = pink, 4 = red, 5 = blue).
- V5** soil, is ordinal and indicates whether the plant grows in dry (1), normal (2), or wet (3) soil.
- V6** preference, is ordinal and gives someone's preference ranking going from 1 to 18.
- V7** height, is interval scaled, the plant's height in centimeters.
- V8** distance, is interval scaled, the distance in centimeters that should be left between the plants.

Source

The reference below.

References

Anja Struyf, Mia Hubert & Peter J. Rousseeuw (1996): Clustering in an Object-Oriented Environment. *Journal of Statistical Software*, **1**. <http://www.stat.ucla.edu/journals/jss/>

gironde

gironde

Description

A list of 4 datasets characterizing conditions of life of 542 cities in Gironde. The four datasets correspond to four thematics relative to conditions of life. Each dataset contains a different number of variables (quantitative and/or qualitative). The first three datasets come from the 2009 population census realized in Gironde by INSEE (Institut National de la Statistique et des Etudes Economiques). The fourth come from an IGN (Institut National de l'Information Geographique et forestiere) database.

Usage

```
data(gironde)
```

Format

A list of 4 data frames.

Value

`gironde$employment`

This data frame contains the description of 542 cities by 9 quantitative variables. These variables are related to employment conditions like, for instance, the average income (income), the percentage of farmers (farmer).

`gironde$housing`

This data frame contains the description of 542 cities by 5 variables (2 qualitative variables and 3 quantitative variables). These variables are related to housing conditions like, for instance, the population density (density), the percentage of council housing within the cities (council).

`gironde$services`

This data frame contains the description of 542 cities by 9 qualitative variables. These variables are related to the number of services within the cities, like, for instance, the number of bakeries (baker) or the number of post office (postoffice).

`gironde$environment`

This data frame contains the description of 542 cities by 4 quantitative variables. These variables are related to the natural environment of the cities, like, for instance the percentage of agricultural land (agricul) or the percentage of buildings (building).

Source

www.INSEE.fr

www.ign.fr

<http://siddt.grenoble.cemagref.fr/>

Multivariate analysis of mixed data: The PCAmixdata R package, M. Chavent, V. Kuentz-Simonet, A. Labenne, J. Saracco, arXiv:1411.4911 [stat.CO]

Lg	<i>Coefficient Lg</i>
----	-----------------------

Description

Compute a data frame with all the Lg coefficients between each matrix of a list

Usage

`Lg(liste.mat)`

Arguments

`liste.mat` a list of G matrix

Value

`Lg` a data frame with G rows and G columns with all the Lg coefficients

Examples

```
V0<-c("a","b","a","a","b")
V01<-c("c","d","e","c","e")
V1<-c(5,4,2,3,6)
V2<-c(8,15,4,6,5)
V3<-c(4,12,5,8,7)
V4<-c("vert","vert","jaune","rouge","jaune")
V5<-c("grand","moyen","moyen","petit","grand")
G1<-data.frame(V0,V01,V1)
```

```
G2<-data.frame(V2,V3)
G3<-data.frame(V4,V5)
liste.mat<-list(G1,G2,G3)
Lg(liste.mat)
```

Lg.pond

Coefficient Lg with ponderation

Description

Compute a data fame with all the Lg coefficients between each matrix of a list ponderated

Usage

```
Lg.pond(liste.mat, ponde)
```

Arguments

liste.mat	a list of G matrix
ponde	a vector of size G with the ponderation associated to each matrix

Value

Lg.pond	a data frame with G rows and G columns with all the Lg coefficients ponderated
---------	--

Examples

```
V0<-c("a","b","a","a","b")
V01<-c("c","d","e","c","e")
V1<-c(5,4,2,3,6)
V2<-c(8,15,4,6,5)
V3<-c(4,12,5,8,7)
V4<-c("vert","vert","jaune","rouge","jaune")
V5<-c("grand","moyen","moyen","petit","grand")
G1<-data.frame(V0,V01,V1)
G2<-data.frame(V2,V3)
G3<-data.frame(V4,V5)
liste.mat<-list(G1,G2,G3)
Lg(liste.mat)
```

Description

Performs multiple factor analysis to analyze a set of individuals (observations) described by several groups of variables. Variables within a group can be a mixture of quantitative and qualitative variables.

Usage

```
MFAmix(data, groups, name.groups, ndim=5, rename.level=FALSE, graph = TRUE,
       axes = c(1, 2))
```

Arguments

data	a data frame with n rows and p columns containing all the variables. This data frame will be split into G groups according to the vector groups.
groups	a vector which gives the groups of the columns in data.
name.groups	a vector of size G which gives the names of the groups.
ndim	number of dimensions kept in the results (by default 5).
rename.level	boolean, if TRUE all the levels of the qualitative variables are renamed as follows: "variable_name=level_name". This prevents to have identical names for the levels.
graph	boolean, if TRUE the following graphics are displayed for the first two dimensions of PCAmix: plot of the individuals coordinates, plot of the squared loadings of variables, plot of the partial axes, plot of the correlation circle (if quantitative variables are available), plot of the levels component map (if qualitative variables are available).
axes	a length 2 vector specifying the axes to plot.

Details

Multiple Factor Analysis (MFA) developed by Escofier and Pages in 1983 is a method of factorial analysis to deal with multiple groups of variables collected on the same observations. The main idea of MFA is to normalize each group by dividing all the variables belonging to this group by the first eigenvalue coming from the Principal Component Analysis (PCA) of this group. Then, a usual PCA on all the weighted variables taken together is applied. Initially this method has been developed for groups only containing quantitative variables. Afterwards this method has been improved to deal simultaneously with groups of qualitative variables and groups of quantitative variables. The MFAmix method allows to perform MFA method for groups containing a mixture of quantitative and qualitative variables

One of the outputs available in the MFAmix method are the squared loadings (sqload). Squared loadings for a qualitative variable are correlation ratios between the variable and the principal components. For a quantitative variable, squared loadings are the squared correlation between the variable and the principal components.

Some others outputs are specific to MFA:

- Coordinates of groups are the sum of the absolute contributions of variables belonging to the groups,
- Partial individuals coordinates are factor coordinates of individuals according to a specific group. The partial coordinates can be achieved by projecting the data set of each group onto the principal component space of MFAmix,
- Partial axes of a group are correlation between each principal components of the separated analyses of the group and the principal components of MFAmix.

Value

<code>eig</code>	a matrix containing the eigenvalues, the percentages of variance and the cumulative percentages of variance.
<code>ind</code>	a list containing the results for the individuals (observations): <ul style="list-style-type: none"> • <code>\$coord</code>: factor coordinates (scores) of the individuals, • <code>\$contrib</code>: absolute contributions of the individuals, • <code>\$contrib.pct</code>: relative contributions of the individuals, • <code>\$cos2</code>: squared cosinus of the individuals.
<code>quanti</code>	a list containing the results for the quantitative variables: <ul style="list-style-type: none"> • <code>\$coord</code>: factor coordinates (scores) of the quantitative variables, • <code>\$contrib</code>: absolute contributions of the quantitative variables, • <code>\$contrib.pct</code>: relative contributions of the quantitative variables (in percentage), • <code>\$cos2</code>: squared cosinus of the quantitative variables.
<code>levels</code>	a list containing the results for the levels of the qualitative variables: <ul style="list-style-type: none"> • <code>\$coord</code>: factor coordinates (scores) of the levels, • <code>\$contrib</code>: absolute contributions of the levels, • <code>\$contrib.pct</code>: relative contributions of the levels (in percentage), • <code>\$cos2</code>: squared cosinus of the levels.
<code>quali</code>	a list containing the results for the qualitative variables: <ul style="list-style-type: none"> • <code>\$contrib</code>: absolute contributions of the qualitative variables (sum of absolute contributions of the levels of the qualitative variable), • <code>\$contrib.pct</code>: relative contributions (in percentage) of the qualitative variables (sum of relative contributions of the levels of the qualitative variable).
<code>sqload</code>	a matrix of dimension (p, ndim) containing the squared loadings of the quantitative and qualitative variables.
<code>coef</code>	the coefficients of the linear combinations used to construct the principal components of MFAmix, and to predict coordinates (scores) of new observations in the function <code>predict.MFAmix</code> .
<code>eig.separate</code>	a matrix containing the ndim first eigenvalues of the separated analyses of each group.
<code>separate.analyses</code>	the results for the separated analyses of each group.

<code>groups</code>	a list containing the results for the groups: <ul style="list-style-type: none"> • <code>\$Lg</code>: Lg coefficients between groups, • <code>\$RV</code>: RV coefficients between groups, • <code>\$contrib</code>: contributions of the groups (sum of variable contributions belonging to the group) • <code>\$contrib.pct</code>: relative contributions of the groups times 100,
<code>partial.axes</code>	a matrix containing the coordinates of the partial axes.
<code>ind.partial</code>	a list of G matrices containing the coordinates of the partial individuals.
<code>listvar.group</code>	list the variables in each group. It is usefull to check the adequacy between the vector groups and the vector <code>name.groups</code> .
<code>global.pca</code>	an object of class PCAmix containing the results of MFAmix considered as a unique PCAmix.

Author(s)

Amaury Labenne <amaury.labenne@irstea.fr>, Marie Chavent, Vanessa Kuentz, Benoit Liquet, Jerome Saracco

References

- Chavent M., Kuentz-Simonet V., Labenne A., Saracco J., Multivariate analysis of mixed data: The PCAmixdata R package, arXiv:1411.4911 [stat.CO].
- Escofier, B. and Pages, J. (1994). Multiple factor analysis (afmult package). Computational statistics & data analysis, 18(1):121-140.
- Le, S., Josse, J., and Husson, F. (2008). Factominer: an r package for multivariate analysis. Journal of statistical software, 25(1):1-18.

See Also

[print.MFAmix](#), [summary.MFAmix](#), [predict.MFAmix](#), [plot.MFAmix](#)

Examples

```
data(gironde)

class.var<-c(rep(1,9),rep(2,5),rep(3,9),rep(4,4))
names <- c("employment","housing","services","environment")

dat<-cbind(gironde$employment[1:20,],gironde$housing[1:20,],
           gironde$services[1:20,],gironde$environment[1:20,])

res<-MFAmix(data=dat,groups=class.var,
            name.groups=names, rename.level=TRUE, ndim=3,graph=FALSE)

summary(res)
```

PCAmix

*Principal component analysis of mixed data***Description**

Performs principal component analysis of a set of individuals (observations) described by a mixture of qualitative and quantitative variables. PCAmix includes the ordinary principal component analysis (PCA) and multiple correspondence analysis (MCA) as special cases.

Usage

```
PCAmix(X.quanti=NULL,X.quali=NULL,ndim=5,rename.level=FALSE,
       weight.col.quanti=NULL,weight.col.quali=NULL,graph=TRUE)
```

Arguments

<code>X.quanti</code>	a numeric matrix of data, or an object that can be coerced to such a matrix (such as a numeric vector or a data frame with all numeric columns).
<code>X.quali</code>	a categorical matrix of data, or an object that can be coerced to such a matrix (such as a character vector, a factor or a data frame with all factor columns).
<code>ndim</code>	number of dimensions kept in the results (by default 5).
<code>rename.level</code>	boolean, if TRUE all the levels of the qualitative variables are renamed as follows: "variable_name=level_name". This prevents to have identical names of the levels.
<code>graph</code>	boolean, if TRUE the following graphics are displayed for the first two dimensions of PCAmix: component map of the individuals, plot of the squared loadings of all the variables (quantitative and qualitative), plot of the correlation circle (if quantitative variables are available), component map of the levels (if qualitative variables are available).
<code>weight.col.quanti</code>	vector of weights for the quantitative variables.
<code>weight.col.quali</code>	vector of the weights for the qualitative variables.

Details

If `X.quali` is not specified (i.e. NULL), only quantitative variables are available and standard PCA is performed. If `X.quanti` is NULL, only qualitative variables are available and standard MCA is performed.

Missing values are replaced by means for quantitative variables and by zeros in the indicator matrix for qualitative variables.

PCAmix performs squared loadings in (sqload). Squared loadings for a qualitative variable are correlation ratios between the variable and the principal components. For a quantitative variable, squared loadings are the squared correlations between the variable and the principal components.

Note that when all the p variables are qualitative, the factor coordinates (scores) of the n observations are equal to the factor coordinates (scores) of standard MCA times square root of p and the eigenvalues are then equal to the usual eigenvalues of MCA times p . When all the variables are quantitative, PCAmix gives exactly the same results as standard PCA.

Value

<code>eig</code>	a matrix containing the eigenvalues, the percentages of variance and the cumulative percentages of variance.
<code>ind</code>	a list containing the results for the individuals (observations): <ul style="list-style-type: none"> • <code>\$coord</code>: factor coordinates (scores) of the individuals, • <code>\$contrib</code>: absolute contributions of the individuals, • <code>\$contrib.pct</code>: relative contributions of the individuals, • <code>\$cos2</code>: squared cosinus of the individuals.
<code>quanti</code>	a list containing the results for the quantitative variables: <ul style="list-style-type: none"> • <code>\$coord</code>: factor coordinates (scores) of the quantitative variables, • <code>\$contrib</code>: absolute contributions of the quantitative variables, • <code>\$contrib.pct</code>: relative contributions of the quantitative variables (in percentage), • <code>\$cos2</code>: squared cosinus of the quantitative variables.
<code>levels</code>	a list containing the results for the levels of the qualitative variables: <ul style="list-style-type: none"> • <code>\$coord</code>: factor coordinates (scores) of the levels, • <code>\$contrib</code>: absolute contributions of the levels, • <code>\$contrib.pct</code>: relative contributions of the levels (in percentage), • <code>\$cos2</code>: squared cosinus of the levels.
<code>quali</code>	a list containing the results for the qualitative variables: <ul style="list-style-type: none"> • <code>\$contrib</code>: absolute contributions of the qualitative variables (sum of absolute contributions of the levels of the qualitative variable), • <code>\$contrib.pct</code>: relative contributions (in percentage) of the qualitative variables (sum of relative contributions of the levels of the qualitative variable).
<code>sqload</code>	a matrix of dimension $(p, ndim)$ containing the squared loadings of the quantitative and qualitative variables.
<code>coef</code>	the coefficients of the linear combinations used to construct the principal components of PCAmix, and to predict coordinates (scores) of new observations in the function <code>predict.PCAmix</code> .
<code>M</code>	the vector of the weights of the columns used in the Generalized Singular Value Decomposition.

Author(s)

Marie Chavent <marie.chavent@u-bordeaux.fr>, Vanessa Kuentz, Amaury Labenne, Benoit Liquet, Jerome Saracco

References

- Chavent M., Kuentz-Simonet V., Labenne A., Saracco J., Multivariate analysis of mixed data: The PCAmixdata R package, arXiv:1411.4911 [stat.CO].
- Chavent, M., Kuentz, V., Saracco, J. (2011), Orthogonal Rotation in PCAMIX. *Advances in Classification and Data Analysis*, Vol. 6, pp. 131-146.
- Kiers, H.A.L., (1991), Simple structure in Component Analysis Techniques for mixtures of qualitative and quantitative variables, *Psychometrika*, 56, 197-212.
- Le, S., Josse, J., and Husson, F. (2008). Factominer: an r package for multivariate analysis. *Journal of statistical software*, 25(1):1-18.
- Pages, J. (2004). Analyse factorielle de donnees mixtes. *Revue de statistique appliquee*, 52(4):93-111.

See Also

[print.PCAmix](#), [summary.PCAmix](#), [predict.PCAmix](#), [plot.PCAmix](#)

Examples

```
#PCAMIX:
data(wine)
str(wine)
X.quanti <- splitmix(wine)$X.quanti
X.quali <- splitmix(wine)$X.quali
pca<-PCAmix(X.quanti[,1:27],X.quali,ndim=4)
pca<-PCAmix(X.quanti[,1:27],X.quali,ndim=4,graph=FALSE)
pca$eig
pca$ind$coord

#PCA:
data(decathlon)
quali<-decathlon[,13]
pca<-PCAmix(decathlon[,1:10])
pca<-PCAmix(decathlon[,1:10], graph=FALSE)
plot(pca,choice="ind",coloring.ind=quali,cex=0.8,
      posleg="topright",main="Scores")
plot(pca, choice="sqload",main="Squared correlations")
plot(pca, choice="cor",main="Correlation circle")
pca$quanti$coord

#MCA
data(flower)
mca <- PCAmix(X.quali=flower[,1:4],rename.level=TRUE)
mca <- PCAmix(X.quali=flower[,1:4],rename.level=TRUE,graph=FALSE)
plot(mca,choice="ind",main="Scores")
plot(mca,choice="sqload",main="Correlation ratios")
plot(mca,choice="levels",main="Levels")
mca$levels$coord

#Missing values
data(vnf)
```

```
PCAmix(X.quali=vnf, rename.level=TRUE)
vnf2<-na.omit(vnf)
PCAmix(X.quali=vnf2, rename.level=TRUE)
```

PCArrot

Varimax rotation in PCAmix

Description

Orthogonal rotation in PCAmix by maximization of the varimax function expressed in terms of PCAmix squared loadings (correlation ratios for qualitative variables and squared correlations for quantitative variables). PCArrot includes the ordinary varimax rotation in Principal Component Analysis (PCA) and a varimax-type rotation in Multiple Correspondence Analysis (MCA) as special cases.

Usage

```
PCArrot(obj, dim, itermax = 100, graph = TRUE)
```

Arguments

obj	an object of class PCAmix.
dim	number of rotated Principal Components.
itermax	maximum number of iterations in the Kaiser's practical optimization algorithm based on successive pairwise planar rotations.
graph	boolean, if TRUE the following graphs are displayed for the first two dimensions after rotation: plot of the individuals (factor coordinates), plot of the variables (squared loadings) plot of the correlation circle (if quantitative variables are available), plot of the levels component map (if qualitative variables are available).

Details

If X.quali is not specified (i.e. NULL) in the previous PCAmix step, only quantitative variables are available and standard varimax rotation in PCA is performed. If X.quanti is NULL, only qualitative variables are available and varimax-type rotation in MCA is performed. Note that p1 is the number of quantitative variables, p2 is the number of qualitative variables and m is the total number of levels of the p2 qualitative variables.

Value

eig	variances of the ndim dimensions after rotation.
ind\$coord	a n by dim quantitative matrix which contains the coordinates (scores) of the n individuals on the dim rotated principal components.
quanti\$coord	a p1 by dim quantitative matrix which contains the coordinates (loadings) of the p1 quantitative variables after rotation. The coordinates of the quantitative variables after rotation are correlations with the rotated principal components.

<code>levels\$coord</code>	a m by dim quantitative matrix which contains the coordinates of the m levels on the dim rotated principal components.
<code>quali\$coord</code>	a p2 by dim quantitative matrix which contains the coordinates of the p2 qualitative variables on the dim rotated principal components. Coordinates of the qualitative variables after rotation are correlation ratio with the rotated principal components.
<code>coef</code>	coefficients of the linear combinations used to construct the rotated principal components of PCAmix.
<code>theta</code>	angle of rotation if dim is equal to 2.
<code>T</code>	matrix of rotation.

Author(s)

Marie Chavent <marie.chavent@u-bordeaux.fr>, Vanessa Kuentz, Benoit Liquet, Jerome Saracco

References

Chavent, M., Kuentz, V., Saracco, J. (2011), Orthogonal Rotation in PCAMIX. Advances in Classification and Data Analysis, Vol. 6, pp. 131-146.

Chavent M., Kuentz-Simonet V., Labenne A., Saracco J., Multivariate analysis of mixed data: The PCAmixdata R package, arXiv:1411.4911 [stat.CO].

Kiers, H.A.L., (1991), Simple structure in Component Analysis Techniques for mixtures of qualitative and quantitative variables, Psychometrika, 56, 197-212.

See Also

[plot.PCAmix](#), [summary.PCAmix](#), [PCAmix](#), [predict.PCAmix](#)

Examples

```
#PCAMIX:
data(wine)
pca<-PCAmix(X.quanti=wine[,c(3:29)],X.quali=wine[,1:2],ndim=4,graph=FALSE)
pca

rot<-PCArrot(pca,3)
rot
rot$eig #percentages of variances after rotation

plot(rot,choice="ind",coloring.ind=wine[,1],
      posleg="bottomleft", main="Rotated scores")
plot(rot,choice="sqload",main="Squared loadings after rotation")
plot(rot,choice="levels",main="Levels after rotation")
plot(rot,choice="cor",main="Correlation circle after rotation")

#PCA:
```

```

data(decathlon)
quali<-decathlon[,13]
pca<-PCAMix(decathlon[,1:10], graph=FALSE)

rot<-PCARot(pca,3)
plot(rot,choice="ind",coloring.ind=quali,cex=0.8,
      posleg="topright",main="Scores after rotation")
plot(rot, choice="sqload", main="Squared correlations after rotation")
plot(rot, choice="cor", main="Correlation circle after rotation")

#MCA
data(flower)
mca <- PCAMix(X.quali=flower[,1:4],rename.level=TRUE,graph=FALSE)

rot<-PCARot(mca,2)
plot(rot,choice="ind",main="Scores after rotation")
plot(rot, choice="sqload", main="Correlation ratios after rotation")
plot(rot, choice="levels", main="Levels after rotation")

```

plot.MFAMix

Graphical outputs of MFAMix

Description

Displays the graphical outputs of MFAMix. Individuals (observations), quantitative variables and levels of the qualitative variables are plotted as points using their factor coordinates (scores) in MFAMix. All the variables (quantitative and qualitative) are plotted on the same graph as points using their squared loadings. The groups of variables are plotted using their contributions to the component coordinates. Partial axes and partial individuals of separated analyses can also be plotted.

Usage

```

## S3 method for class 'MFAMix'
plot(x, axes = c(1, 2), choice = "axes",
     label=TRUE, coloring.var = "not", coloring.ind=NULL,
     nb.partial.axes=3,col.ind=NULL, col.groups=NULL, partial = NULL,
     lim.cos2.plot = 0, lim.contrib.plot=0, xlim = NULL, ylim = NULL,
     cex = 1, main = NULL, new.plot = FALSE,leg=TRUE,
     posleg="topleft",cex.leg=0.8,
     col.groups.sup=NULL,posleg.sup="topright",nb.paxes.sup=3,...)

```

Arguments

x	an object of class MFAMix.
axes	a length 2 vector specifying the components to plot.

choice	the graph to plot: <ul style="list-style-type: none"> • "ind" for the individuals, • "cor" for the correlation circle of the quantitative variables, • "levels" for the levels of of the qualitative variables, • "sqload" for the plot of the squared loadings of all the variables, • "groups" for the plot of the contributions of the groups, • "axes" for the correlation circle of the partial axes.
label	boolean, if FALSE the labels of the points are not plotted.
coloring.ind	a qualitative variable such as a character vector or a factor of size n (the number of individuals). The individuals are colored according to the levels of this variable. If NULL, the individuals are not colored.
nb.partial.axes	if choice="axes", the maximum number of partial axes related to each group to plot on the correlation circle. By default equal to 3.
col.ind	a vector of colors, of size the number of levels of coloring.ind. If NULL, colors are chosen automatically.
coloring.var	a value to choose among: <ul style="list-style-type: none"> • "type": the variables in the plot of the squared loadings are colored according to their type (quantitative or qualitative), • "groups": the variables are colored according to their group. • NULL: variables are not colored.
col.groups	a vector of colors, of size the number of groups. If NULL, colors are chosen automatically.
partial	a vector of class character with the row names of the individuals, for which the partial individuals should be drawn. By default partial = NULL and no partial points are drawn. Partial points are colored according to col.groups
lim.cos2.plot	a value between 0 and 1. Points with squared cosinus below this value are not plotted.
lim.contrib.plot	a value between 0 and 100. Points with relative contributions (in percentage) below this value are not plotted.
posleg	position of the legend.
xlim	a numeric vectors of length 2, giving the x coordinates range. If NULL (by default) the range is defined automatically (recommended).
ylim	a numeric vectors of length 2, giving the y coordinates range. If NULL (by default) the range is defined automatically (recommended).
main	a string corresponding to the title of the graph to draw.
cex	cf. function par in the graphics package.
new.plot	boolean, if TRUE, a new graphical device is created.
leg	boolean, if TRUE, a legend is displayed.
cex.leg	a numerical value giving the amount by which the legend should be magnified. Default is 0.8.

posleg.sup	position of the legend for the supplementary groups.
col.groups.sup	a vector of colors, of size the number of supplementary groups. If NULL, colors are chosen automatically.
nb.paxes.sup	if choice="axes", the maximum number of partial axes of supplementary groups plotted on the correlation circle. By default equal to 3.
...	arguments to be passed to methods, such as graphical parameters.

Details

The observations can be colored according to the levels of a qualitative variable. The observations, the quantitative variables and the levels can be selected according to their squared cosine (lim.cos2.plot) or their relative contribution (lim.contrib.plot) to the component map. Only points with squared cosine or relative contribution greater than a given threshold are plotted. Note that the relative contribution of a point to the component map (a plan) is the sum of the absolute contributions to each dimension, divided by the sum of the corresponding eigenvalues.

Author(s)

Amaury Labenne, Marie Chavent <marie.chavent@u-bordeaux.fr>, Vanessa Kuentz, Benoit Liquet, Jerome Saracco

References

Chavent M., Kuentz-Simonet V., Labenne A., Saracco J., Multivariate analysis of mixed data: The PAMixdata R package, arXiv:1411.4911 [stat.CO].

Escofier, B. and Pages, J. (1994). Multiple factor analysis (afmult package). Computational statistics & data analysis, 18(1):121-140.

Le, S., Josse, J., and Husson, F. (2008). Factominer: an r package for multivariate analysis. Journal of statistical software, 25(1):1-18.

Examples

```
data(gironde)

class.var<-c(rep(1,9),rep(2,5),rep(3,9),rep(4,4))
names <- c("employment","housing","services","environment")
dat<-cbind(gironde$employment[1:20,],gironde$housing[1:20,],
           gironde$services[1:20,],gironde$environment[1:20,])
res<-MFAMix(data=dat,groups=class.var,
            name.groups=names, rename.level=TRUE, ndim=3,graph=FALSE)

#---- quantitative variables
plot(res,choice="cor",cex=0.6)
plot(res,choice="cor",cex=0.6,coloring.var="groups")
plot(res,choice="cor",cex=0.6,coloring.var="groups",
     col.groups=c("red","yellow","pink","brown"),leg=TRUE)

#----partial axes
plot(res,choice="axes",cex=0.6)
```

```

plot(res,choice="axes",cex=0.6,coloring.var="groups")
plot(res,choice="axes",cex=0.6,coloring.var="groups",
      col.groups=c("red","yellow","pink","brown"),leg=TRUE)

#----groups
plot(res,choice="groups",cex=0.6) #no colors for groups
plot(res,choice="groups",cex=0.6,coloring.var="groups")
plot(res,choice="groups",cex=0.6,coloring.var="groups",
      col.groups=c("red","yellow","pink","blue"))

#----squared loadings
plot(res,choice="sqload",cex=0.8) #no colors for groups
plot(res,choice="sqload",cex=0.8,coloring.var="groups",
      posleg="topright")
plot(res,choice="sqload",cex=0.6,coloring.var="groups",
      col.groups=c("red","yellow","pink","blue"),ylim=c(0,1))

plot(res,choice="sqload",cex=0.8,coloring.var="type",
      cex.leg=0.9,posleg="topright")

#----individuals
plot(res,choice="ind",cex=0.6)

#----individuals with squared cosine greater than 0.5
plot(res,choice="ind",cex=0.6,lim.cos2.plot=0.5)

#----individuals colored with a qualitative variable
nbchem <- gironde$services$chemist[1:20]
plot(res,choice="ind",cex=0.6,coloring.ind=nbchem,
      posleg="topright")
plot(res,choice="ind",coloring.ind=nbchem,
      col.ind=c("pink","brown","darkblue"),label=FALSE,posleg="topright")

#----partial individuals colored by groups
plot(res,choice="ind",partial=c("AUBIAC","ARCACHON"),
      cex=0.6,posleg="bottomright")

#----levels of qualitative variables
plot(res,choice="levels",cex=0.8)
plot(res,choice="levels",cex=0.8,coloring.var="groups")

#levels with squared cosine greater than 0.6
plot(res,choice="levels",cex=0.8, lim.cos2.plot=0.6)

#supplementary groups
data(wine)
X.quant1 <- splitmix(wine)$X.quant1[,1:5]
X.qual1 <- splitmix(wine)$X.qual1[,1,drop=FALSE]
X.quant1.sup <- splitmix(wine)$X.quant1[,28:29]
X.qual1.sup <- splitmix(wine)$X.qual1[,2,drop=FALSE]
data <- cbind(X.quant1,X.qual1)
data.sup <- cbind(X.quant1.sup,X.qual1.sup)

```

```

groups <-c(1,2,2,3,3,1)
name.groups <- c("G1","G2","G3")
groups.sup <- c(1,1,2)
name.groups.sup <- c("Gsup1","Gsup2")
mfa <- MFAMix(data,groups,name.groups,ndim=4,rename.level=TRUE,graph=FALSE)
mfa.sup <- supvar(mfa,data.sup,groups.sup,name.groups.sup,rename.level=TRUE)
plot(mfa.sup,choice="sqload",coloring.var="groups")
plot(mfa.sup,choice="axes",coloring.var="groups")
plot(mfa.sup,choice="groups",coloring.var="groups")
plot(mfa.sup,choice="levels",coloring.var="groups")
plot(mfa.sup,choice="levels")
plot(mfa.sup,choice="cor",coloring.var = "groups")

```

plot.PCAmix

Graphical outputs of PCAmix and PCARot

Description

Displays the graphical outputs of PCAmix and PCARot. The individuals (observations), the quantitative variables and the levels of the qualitative variables are plotted as points using their factor coordinates (scores). All the variables (quantitative and qualitative) are plotted as points on the same graph using their squared loadings.

Usage

```

## S3 method for class 'PCAmix'
plot(x,axes = c(1, 2), choice = "ind",label=TRUE,
      coloring.ind=NULL,col.ind=NULL, coloring.var=FALSE,
      lim.cos2.plot=0,lim.contrib.plot=0, posleg="topleft",
      xlim=NULL,ylim=NULL, cex=1,leg=TRUE,main=NULL,cex.leg=1, ...)

```

Arguments

x	an object of class PCAmix obtained with the function PCAmix or PCARot.
axes	a length 2 vector specifying the components to plot.
choice	the graph to plot: <ul style="list-style-type: none"> • "ind" for the individuals component map, • "cor" for the correlation circle if quantitative variables are available in the data, • "levels" for the levels components map (if qualitative variables are available in the data), • "sqload" for the plot of the squared loadings of all the variables.
label	boolean, if FALSE the labels of the points are not plotted.
coloring.ind	a qualitative variable such as a character vector or a factor of size n (the number of individuals). The individuals are colored according to the levels of this variable. If NULL, the individuals are not colored.

<code>col.ind</code>	a vector of colors, of size the number of levels of <code>coloring.ind</code> . If NULL, colors are chosen automatically.
<code>coloring.var</code>	boolean, if TRUE, the variables in the plot of the squared loadings are colored according to their type (quantitative or qualitative).
<code>lim.cos2.plot</code>	a value between 0 and 1. Points with squared cosinus below this value are not plotted.
<code>lim.contrib.plot</code>	a value between 0 and 100. Points with relative contributions (in percentage) below this value are not plotted.
<code>posleg</code>	position of the legend.
<code>xlim</code>	a numeric vectors of length 2, giving the x coordinates range. If NULL (by default) the range is defined automatically (recommended).
<code>ylim</code>	a numeric vectors of length 2, giving the y coordinates range. If NULL (by default) the range is defined automatically (recommended).
<code>main</code>	a string corresponding to the title of the graph to draw.
<code>cex</code>	cf. function <code>par</code> in the graphics package
<code>leg</code>	if TRUE a legend is displayed for <code>choice="sqload"</code> and <code>coloring.var="type"</code> .
<code>cex.leg</code>	a numerical value giving the amount by which the legend should be magnified. Default is 0.8.
<code>...</code>	arguments to be passed to methods, such as graphical parameters.

Details

The observations can be colored according to the levels of a qualitative variable. The observations, the quantitative variables and the levels can be selected according to their squared cosine (`lim.cos2.plot`) or their relative contribution (`lim.contrib.plot`) to the component map. Only points with squared cosine or relative contribution greater than a given threshold are plotted. Note that the relative contribution of a point to the component map (a plan) is the sum of the absolute contributions to each dimension, divided by the sum of the corresponding eigenvalues.

Author(s)

Amaury Labenne <amaury.labenne@irstea.fr>, Marie Chavent, Vanessa Kuentz, Benoit Liquet, Jerome Saracco

References

Chavent M., Kuentz-Simonet V., Labenne A., Saracco J., Multivariate analysis of mixed data: The PCAmixdata R package, arXiv:1411.4911 [stat.CO].

Le, S., Josse, J., and Husson, F. (2008). Factominer: an r package for multivariate analysis. Journal of statistical software, 25(1):1-18.

See Also

[summary.PCAmix](#), [PCAmix](#), [PCArrot](#),

Examples

```
data(gironde)
base <- gironde$housing[1:20,]
X.quanti <- splitmix(base)$X.quanti
X.quali <- splitmix(base)$X.quali
res<-PCAmix(X.quanti, X.quali, rename.level=TRUE, ndim=3, graph=FALSE)

#----quantitative variables on the correlation circle
plot(res,choice="cor",cex=0.8)

#----individuals component map
plot(res,choice="ind",cex=0.8)

#----individuals colored with the qualitative variable "houses"
houses <- X.quali$houses
plot(res,choice="ind",cex=0.6,coloring.ind=houses)

#----individuals selected according to their cos2
plot(res,choice="ind",cex=0.6,lim.cos2.plot=0.8)

#----all the variables plotted with the squared loadings
plot(res,choice="sqload",cex=0.8)

#----variables colored according to their type (quanti or quali)
plot(res,choice="sqload",cex=0.8,coloring.var=TRUE)

#----levels component map
plot(res,choice="levels",cex=0.8)

#----example with supplementary variables
data(wine)
X.quanti <- splitmix(wine)$X.quanti[,1:5]
X.quali <- splitmix(wine)$X.quali[,1,drop=FALSE]
X.quanti.sup <- splitmix(wine)$X.quanti[,28:29]
X.quali.sup <- splitmix(wine)$X.quali[,2,drop=FALSE]
pca<-PCAmix(X.quanti,X.quali,ndim=4,graph=FALSE)
pca2 <- supvar(pca,X.quanti.sup,X.quali.sup)
plot(pca2,choice="levels")
plot(pca2,choice="cor")
plot(pca2,choice="sqload")
```

predict.MFamix

Prediction of new scores in MFamix

Description

This function performs the scores of new observations on the principal components of MFamix. In other words, this function is projecting the new observations onto the principal components of MFamix obtained previously on a separated dataset. Note that the new observations must be

described with the same variables than those used in MFamix. The groups of variables must also be identical.

Usage

```
## S3 method for class 'MFamix'
predict(object, data, rename.level=FALSE,...)
```

Arguments

<code>object</code>	an object of class MFamix (output of the function MFamix)
<code>data</code>	a data frame containing the description of the new observations on all the variables. This data frame will be split into G groups according to the vector groups.
<code>rename.level</code>	boolean, if TRUE all the levels of the qualitative variables are renamed as follows: "variable_name=level_name". This prevents to have identical names for the levels.
<code>...</code>	further arguments passed to or from other methods. They are ignored in this function.

Value

Returns the matrix of the scores of the new observations on the principal components or on the rotated principal components of MFamix.

Author(s)

Amaury Labenne <amaury.labenne@irstea.fr>, Marie Chavent, Vanessa Kuentz, Benoit Liquet, Jerome Saracco

References

Chavent M., Kuentz-Simonet V., Labenne A., Saracco J., Multivariate analysis of mixed data: The PCAmixdata R package, arXiv:1411.4911 [stat.CO].

See Also

[MFamix](#)

Examples

```
data(gironde)
class.var<-c(rep(1,9),rep(2,5),rep(3,9),rep(4,4))
names<-c("employment","housing","services","environment")
dat<-cbind(gironde$employment,gironde$housing,
           gironde$services,gironde$environment)
n <- nrow(dat)
set.seed(10)
sub <- sample(1:n,520)

res<-MFamix(data=dat[sub,],groups=class.var,
```

```

name.groups=names, rename.level=TRUE,
ndim=3,graph=FALSE)

#Predict scores of new data
pred<-predict(res,data=dat[-sub,])
plot(res,choice="ind",cex=0.6,lim.cos2.plot=0.7)
points(pred[1:5,c(1,2)],col=2,pch=16,cex=0.6)
text(pred[1:5,c(1,2)], labels = rownames(dat[-sub,])[1:5],
      col=2,pos=3,cex=0.6)

```

predict.PCAmix

Prediction of new scores in PCAmix or PCArot

Description

This function performs the scores of new observations on the principal components of PCAmix. If the components have been rotated, this function performs the scores of the new observations on the rotated principal components. In other words, this function is projecting the new observations onto the principal components of PCAmix (or PCArot) obtained previously on a separated dataset. Note that the new observations must be described with the same variables than those used in PCAmix (or PCArot).

Usage

```

## S3 method for class 'PCAmix'
predict(object, X.quanti = NULL, X.quali = NULL, rename.level=FALSE,...)

```

Arguments

object	an object of class PCAmix (output of the function PCAmix or PCArot).
X.quanti	a numeric data matrix or an object that can be coerced to such a matrix (such as a numeric vector or a data frame with all numeric columns)
X.quali	a categorical matrix of data, or an object that can be coerced to such a matrix (such as a character vector, a factor or a data frame with all factor columns)
rename.level	boolean, if TRUE all the levels of the qualitative variables are renamed as follows: "variable_name=level_name". This prevents to have identical names for the levels.
...	further arguments passed to or from other methods. They are ignored in this function.

Value

Returns the matrix of the scores of the new observations on the principal components or on the rotated principal components of PCAmix.

Author(s)

Marie Chavent <marie.chavent@math.u-bordeaux1.fr>, Vanessa Kuentz, Benoit Liquet, Jerome Saracco

References

Chavent M., Kuentz-Simonet V., Labenne A., Saracco J., Multivariate analysis of mixed data: The PCAMixdata R package, arXiv:1411.4911 [stat.CO].

See Also

[PCAMix](#)

Examples

```
# quantitative data
data(decathlon)
n <- nrow(decathlon)
sub <- sample(1:n,20)
pca<-PCAMix(decathlon[sub,1:10], graph=FALSE)
predict(pca,decathlon[-sub,1:10])
rot <- PCARot(pca,dim=4)
predict(rot,decathlon[-sub,1:10])

# quantitative and qualitative data
data(wine)
str(wine)
X.quanti <- splitmix(wine)$X.quanti
X.quali <- splitmix(wine)$X.quali
pca<-PCAMix(X.quanti[,1:27],X.quali,ndim=4,graph=FALSE)
n <- nrow(wine)
sub <- sample(1:n,10)
pca<-PCAMix(X.quanti[sub,1:27],X.quali[sub,],ndim=4)
pred <- predict(pca,X.quanti[-sub,1:27],X.quali[-sub,])
plot(pca,axes=c(1,2))
points(pred[,c(1,2)],col=2,pch=16)
text(pred[,c(1,2)], labels = rownames(X.quanti[-sub,1:27]), col=2,pos=3)
```

`print.MFAMix`

Print a 'MFAMix' object

Description

This is a method for the function `print` for objects of the class `MFAMix`.

Usage

```
## S3 method for class 'MFAMix'
print(x, ...)
```


Arguments

`x` an object of class MFAMix generated by the function [PCAMix](#).
`...` further arguments to be passed to or from other methods. They are ignored in this function.

See Also

[MFAMix](#)

print.PCAmix	<i>Print a 'PCAMix' object</i>
--------------	--------------------------------

Description

This is a method for the function print for objects of the class PCAMix.

Usage

```
## S3 method for class 'PCAMix'
print(x, ...)
```

Arguments

`x` an object of class PCAMix generated by the functions [PCAMix](#) and [PCArrot](#).
`...` further arguments to be passed to or from other methods. They are ignored in this function.

See Also

[PCAMix](#), [PCArrot](#)

protein	<i>Protein data</i>
---------	---------------------

Description

The data measure the amount of protein consumed for nine food groups in 25 European countries. The nine food groups are red meat (RedMeat), white meat (WhiteMeat), eggs (Eggs), milk (Milk), fish (Fish), cereal (Cereal), starch (Starch), nuts (Nuts), and fruits and vegetables (FruitVeg).

Format

A data frame with 25 rows (the European countries) and 9 columns (the food groups)

Source

Originated by A. Weber and cited in Hand et al., A Handbook of Small Data Sets, (1994, p. 297).

 recod

Recoding of the data matrices

Description

Recoding of the quantitative and of the qualitative data matrix.

Usage

```
recod(X.quanti, X.quali, rename.level=FALSE)
```

Arguments

<code>X.quanti</code>	a numerical data matrix.
<code>X.quali</code>	a categorical data matrix.
<code>rename.level</code>	boolean, if TRUE all the levels of the qualitative variables are renamed as follows: "variable_name=level_name".

Value

<code>X</code>	<code>X.quanti</code> and <code>X.quali</code> concatenated in a single matrix.
<code>Y</code>	<code>X.quanti</code> with missing values replaced with mean values concatenated with the indicator matrix of <code>X.quali</code> with missing values replaced by zeros.
<code>Z</code>	<code>X.quanti</code> standardized (centered and reduced by standard deviations) concatenated with the indicator matrix of <code>X.quali</code> centered and reduced with the square roots of the relative frequencies of the categories.
<code>W</code>	<code>X.quanti</code> standardized (centered and reduced by standard deviations) concatenated with the indicator matrix of <code>X.quali</code> centered.
<code>n</code>	the number of observations.
<code>p</code>	the total number of variables
<code>p1</code>	the number of quantitative variables
<code>p2</code>	the number of qualitative variables
<code>g</code>	the means of the columns of <code>Y</code>
<code>s</code>	the standard deviations of the columns of <code>Y</code>
<code>G</code>	The indicator matrix of <code>X.quali</code> with missing values replaced by 0
<code>Gcod</code>	The indicator matrix <code>G</code> reduced with the square roots of the relative frequencies of the categories

recodqual	<i>Recoding of the qualitative data matrix.</i>
-----------	---

Description

Recoding of the qualitative data matrix.

Usage

```
recodqual(X, rename.level=FALSE)
```

Arguments

X	the qualitative data matrix.
rename.level	boolean, if TRUE all the levels of the qualitative variables are renamed as follows: "variable_name=level_name".

Value

G	The indicator matrix of X with missing values replaced by 0.
---	--

Examples

```
data(vnf)
X <- vnf[1:10,9:12]
tab.disjonctif.NA(X)
recodqual(X)
```

recodquant	<i>Recoding of the quantitative data matrix</i>
------------	---

Description

Recoding of the quantitative data matrix.

Usage

```
recodquant(X)
```

Arguments

X	the quantitative data matrix.
---	-------------------------------

Value

Z	the standardized quantitative data matrix (centered and reduced with the standard deviations.)
g	the means of the columns of X
s	the standard deviations of the columns of X (population version with 1/n)
Xcod	The quantitative matrix X with missing values replaced with the column mean values.

Examples

```
data(decathlon)
X <- decathlon[1:5,1:5]
X[1,2] <- NA
X[2,3] <-NA
rec <- recodquant(X)
```

RV	<i>Coefficient RV</i>
----	-----------------------

Description

Computes RV coefficients between matrices.

Usage

```
RV(liste.mat)
```

Arguments

liste.mat	a list of k matrices.
-----------	-----------------------

Value

RV	a (k, k) matrix containing the RV coefficients.
----	---

References

Escofier B et Pages J (1998), Analyses factorielles simples et multiples, Dunod, 3e ed.
Pages J, Analyse factorielle multiple appliquee aux variables qualitatives et aux donnees mixtes.
Revue de statistique appliquee, tome 50, num 4 (2002), p. 5-37

See Also

[RV.pond](#), [Lg](#), [Lg.pond](#),

Examples

```
V0<-c("a","b","a","a","b")
V01<-c("c","d","e","c","e")
V1<-c(5,4,2,3,6)
V2<-c(8,15,4,6,5)
V3<-c(4,12,5,8,7)
V4<-c("vert","vert","jaune","rouge","jaune")
V5<-c("grand","moyen","moyen","petit","grand")
G1<-data.frame(V0,V01,V1)
G2<-data.frame(V2,V3)
G3<-data.frame(V4,V5)
liste.mat<-list(G1,G2,G3)
RV(liste.mat)
```

RV.pond

*Coefficient RV with ponderation***Description**

Computes weighted RV coefficients between matrices.

Usage

```
RV.pond(liste.mat, ponde)
```

Arguments

liste.mat	a list of k matrices.
ponde	vector of size k with the ponderation of each matrix.

Value

Returns a matrix of dimension (k, k) with all the weighted RV coefficients.

References

Escofier B et Pages J (1998), Analyses factorielles simples et multiples, Dunod, 3e ed.
 Pages J, Analyse factorielle multiple appliquee aux variables qualitatives et aux donnees mixtes.
 Revue de statistique appliquee, tome 50, num 4 (2002), p. 5-37

See Also

[RV](#), [Lg](#), [Lg.pond](#),

Examples

```

V0<-c("a","b","a","a","b")
V01<-c("c","d","e","c","e")
V1<-c(5,4,2,3,6)
V2<-c(8,15,4,6,5)
V3<-c(4,12,5,8,7)
V4<-c("vert","vert","jaune","rouge","jaune")
V5<-c("grand","moyen","moyen","petit","grand")
G1<-data.frame(V0,V01,V1)
G2<-data.frame(V2,V3)
G3<-data.frame(V4,V5)
liste.mat<-list(G1,G2,G3)
ponderation<-c(1,2,1)
RV.pond(liste.mat,ponderation)

```

sol.2dim

*Varimax rotation for PCAmix in two dimensions***Description**

Varimax rotation for PCAmix in two dimensions

Usage

```
sol.2dim(A, indexj, p, p1)
```

Arguments

A	matrix of loadings
indexj	a vector with the number of the variable associated with each row of A
p	the total number of variables
p1	the number of quantitative variables

Value

theta	the angle of rotation
T	the matrix of rotation

splitgroups	<i>splitgroups</i>
-------------	--------------------

Description

If the p variables of a data matrix of dimension (n,p) are separated into G groups, this functions splits this data matrix into G datasets according the groups membership.

Usage

```
splitgroups(data, groups, name.groups)
```

Arguments

data	the a data matrix into G datasets with n rows and p columns.
groups	a vector of size p whose values indicate at which group belongs each variable.
name.groups	a vector of size G which contains names for each group we want to create. The names are given in the increasing order of the numbers of the groups.

Value

data.groups	a list of G data matrix: one matrix for each group.
listvar.groups	The list of the variables in each group.

Examples

```
data(decathlon)
split <- splitgroups(decathlon,groups=c(rep(1,10),2,2,3),
  name.groups=c("Epreuve","Classement","Competition"))
split$data.groups$Epreuve
```

splitmix	<i>splitmix</i>
----------	-----------------

Description

Splits a mixed data matrix in two data sets: one with the quantitative variables and one with the qualitative variables. Here, the columns of class "integer" are considered quantitative. If you want this column to be considered as qualitative, it must be of class character or factor.

Usage

```
splitmix(data)
```

Arguments

data a data matrix or a data.frame with a mixture of quantitative and qualitative variables.

Value

X.quant a data matrix containing only the quantitative variables.
 X.quali A data.frame containing only the qualitative variables.

Examples

```
data(decathlon)
data.split <- splitmix(decathlon)
data.split$X.quant
data.split$X.quali
```

summary.MFamix	<i>Summary of a 'MFamix' object</i>
----------------	-------------------------------------

Description

This is a method for the function summary for objects of the class MFamix.

Usage

```
## S3 method for class 'MFamix'
summary(object, ...)
```

Arguments

object an object of class MFamix obtained with the function MFamix.
 ... further arguments passed to or from other methods.

Value

Returns the total number of observations, the number of quantitative variables, the number of qualitative variables with the total number of levels. And all those values are also given by groups.

See Also

[plot.MFamix,MFamix](#)

summary.PCAmix	<i>Summary of a 'PCAmix' object</i>
----------------	-------------------------------------

Description

This is a method for the function summary for objects of the class PCAmix.

Usage

```
## S3 method for class 'PCAmix'  
summary(object, ...)
```

Arguments

object	an object of class PCAmix obtained with the function PCAmix or PCArot.
...	further arguments passed to or from other methods.

Value

Returns the matrix of squared loadings. For quantitative variables (resp. qualitative), squared loadings are the squared correlations (resp. the correlation ratios) with the scores or with the rotated (standardized) scores.

See Also

[plot.PCAmix, PCAmix, PCArot,](#)

Examples

```
data(wine)  
X.quanti <- wine[,c(3:29)]  
X.quali <- wine[,c(1,2)]  
pca<-PCAmix(X.quanti,X.quali,ndim=4, graph=FALSE)  
summary(pca)  
  
rot<-PCArot(pca,3,graph=FALSE)  
summary(rot)
```

supvar	<i>Supplementary variables projection</i>
--------	---

Description

supvar is a generic function for adding supplementary variables in PCAMix or MFAMix. The function invokes two methods which depend on the class of the first argument.

Usage

```
supvar(obj, ...)
```

Arguments

obj	an object of class PCAMix or MFAMix.
...	further arguments passed to or from other methods.

Details

This generic function has two methods [supvar.PCAMix](#) and [supvar.MFAMix](#)

supvar.MFAMix	<i>Supplementary variables in MFAMix</i>
---------------	--

Description

Performs the coordinates of supplementary variables and groups on the component of an object of class MFAMix.

Usage

```
## S3 method for class 'MFAMix'
supvar(obj, data.sup, groups.sup, name.groups.sup,
       rename.level = FALSE, ...)
```

Arguments

obj	an object of class MFAMix.
data.sup	a numeric matrix of data.
groups.sup	a vector which gives the groups of the columns in data.sup.
name.groups.sup	a vector which gives the names of the supplementary groups.
rename.level	boolean, if TRUE all the levels of the qualitative variables are renamed as follows: "variable_name=level_name". This prevents to have identical names of the levels.
...	further arguments passed to or from other methods.

Examples

```
data(wine)
X.quanti <- splitmix(wine)$X.quanti[,1:5]
X.quali <- splitmix(wine)$X.quali[,1,drop=FALSE]
X.quanti.sup <- splitmix(wine)$X.quanti[,28:29]
X.quali.sup <- splitmix(wine)$X.quali[,2,drop=FALSE]
data <- cbind(X.quanti,X.quali)
data.sup <- cbind(X.quanti.sup,X.quali.sup)
groups <- c(1,2,2,3,3,1)
name.groups <- c("G1","G2","G3")
groups.sup <- c(1,1,2)
name.groups.sup <- c("Gsup1","Gsup2")
mfa <- MFAmix(data,groups,name.groups,ndim=4,rename.level=TRUE,graph=FALSE)
mfa.sup <- supvar(mfa,data.sup,groups.sup,name.groups.sup,rename.level=TRUE)
```

supvar.PCAmix

Supplementary variables in PCAmix

Description

Performs the coordinates of supplementary variables on the component of an object of class PCAmix.

Usage

```
## S3 method for class 'PCAmix'
supvar(obj, X.quanti.sup = NULL, X.quali.sup = NULL,
       rename.level = FALSE, ...)
```

Arguments

obj	an object of class PCAmix.
X.quanti.sup	a numeric matrix of data.
X.quali.sup	a categorical matrix of data.
rename.level	boolean, if TRUE all the levels of the qualitative variables are renamed as follows: "variable_name=level_name". This prevents to have identical names of the levels.
...	further arguments passed to or from other methods.

See Also

[PCAmix](#)

Examples

```
data(wine)
X.quanti <- splitmix(wine)$X.quanti[,1:5]
X.quali <- splitmix(wine)$X.quali[,1,drop=FALSE]
X.quanti.sup <-splitmix(wine)$X.quanti[,28:29]
X.quali.sup <-splitmix(wine)$X.quali[,2,drop=FALSE]
pca<-PCAmix(X.quanti,X.quali,ndim=4,graph=FALSE)
pcasup <- supvar(pca,X.quanti.sup,X.quali.sup)
```

svd.triplet

Singular Value Decomposition of a Matrix

Description

Compute the singular-value decomposition of a rectangular matrix with weights for rows and columns. Borrowed from the 'FactoMineR' package and, used as internal function

Usage

```
svd.triplet(X, row.w=NULL, col.w=NULL, ncp=Inf)
```

Arguments

X	a data matrix
row.w	vector with the weights of each row (NULL by default and the weights are uniform)
col.w	vector with the weights of each column (NULL by default and the weights are uniform)
ncp	the number of components kept for the outputs

Details

This function has been taken from the package **FactoMineR**. It is then identical.

Value

vs	a vector containing the singular values of 'x';
u	a matrix whose columns contain the left singular vectors of 'x';
v	a matrix whose columns contain the right singular vectors of 'x'.

See Also

[svd](#)

tab.disjonctif.NA	<i>Built an indicator matrix</i>
-------------------	----------------------------------

Description

This function built the indicator matrix of a qualitative data matrix. Missing observations are indicated as NA.

Usage

```
tab.disjonctif.NA(tab, rename.level=FALSE)
```

Arguments

tab	a categorical data matrix.
rename.level	boolean, if TRUE all the levels of the qualitative variables are renamed as follows: "variable_name=level_name".

Details

This function uses the code of the function "tab.disjonctif" implemented in the package **FactoMineR** but is different. Here, a NA value appears when a category has not been observed in a row. In the function "tab.disjonctif" of the package **FactoMineR**, a new column is created in that case. In the output of the function recodqual, the NA are replaced with 0.

Value

Returns the indicator matrix with NA for missing observations.

Examples

```
data(vnf)
X <- vnf[1:10, 9:12]
tab.disjonctif.NA(X)
recodqual(X)
```

vnf	<i>User satisfaction survey with 1232 individuals and 14 questions</i>
-----	--

Description

A user satisfaction survey of pleasure craft operators on the “Canal des Deux Mers”, located in South of France, was carried out by the public corporation “Voies Navigables de France” (VNF) responsible for managing and developing the largest network of navigable waterways in Europe

Usage

```
data(vnf)
```

Format

A data frame with 1232 observations and 14 qualitative variables.

Source

Josse, J., Chavent, M., Liquet, B. and Husson, F. (2012). Handling missing values with Regularized Iterative Multiple Correspondence Analysis. *Journal of classification*, Vol. 29, pp. 91-116.

wine	<i>Wine</i>
------	-------------

Description

The data used here refer to 21 wines of Val de Loire.

Usage

```
data(wine)
```

Format

A data frame with 21 rows (the number of wines) and 31 columns: the first column corresponds to the label of origin, the second column corresponds to the soil, and the others correspond to sensory descriptors.

Source

Centre de recherche INRA d'Angers

Le, S., Josse, J. & Husson, F. (2008). FactoMineR: An R Package for Multivariate Analysis. *Journal of Statistical Software*. 25(1). pp. 1-18.

Index

- *Topic **algebra**
 - svd.triplet, [36](#)
- *Topic **datasets**
 - decathlon, [2](#)
 - flower, [3](#)
 - gironde, [4](#)
 - vnf, [37](#)
 - wine, [38](#)
- *Topic **multivariate, mixed type data, data structured in groups**
 - MFAmix, [7](#)
- *Topic **multivariate**
 - PCARot, [13](#)
- *Topic **principal component analysis, mixed type data**
 - PCAmix, [10](#)
- *Topic **print**
 - print.MFAmix, [24](#)
 - print.PCAmix, [25](#)
- decathlon, [2](#)
- dogs, [3](#)
- flower, [3](#)
- gironde, [4](#)
- Lg, [5](#), [28](#), [29](#)
- Lg.pond, [6](#), [28](#), [29](#)
- MFAmix, [7](#), [22](#), [25](#), [32](#)
- PCAmix, [10](#), [14](#), [20](#), [24](#), [25](#), [33](#), [35](#)
- PCARot, [13](#), [20](#), [25](#), [33](#)
- plot.MFAmix, [9](#), [15](#), [32](#)
- plot.PCAmix, [12](#), [14](#), [19](#), [33](#)
- predict.MFAmix, [8](#), [9](#), [21](#)
- predict.PCAmix, [11](#), [12](#), [14](#), [23](#)
- print.MFAmix, [9](#), [24](#)
- print.PCAmix, [12](#), [25](#)
- protein, [25](#)
- recod, [26](#)
- recodqual, [27](#)
- recodquant, [27](#)
- RV, [28](#), [29](#)
- RV.pond, [28](#), [29](#)
- sol.2dim, [30](#)
- splitgroups, [31](#)
- splitmix, [31](#)
- summary.MFAmix, [9](#), [32](#)
- summary.PCAmix, [12](#), [14](#), [20](#), [33](#)
- supvar, [34](#)
- supvar.MFAmix, [34](#), [34](#)
- supvar.PCAmix, [34](#), [35](#)
- svd, [36](#)
- svd.triplet, [36](#)
- tab.disjonctif.NA, [37](#)
- vnf, [37](#)
- wine, [38](#)