Package 'ClustOfVar'

April 23, 2025

Type Package

Title Clustering of Variables

Version 1.2

Author Marie Chavent [aut, cre],
Vanessa Kuentz [aut],
Benoit Liquet [aut],
Jerome Saracco [aut]

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

Description
Cluster analysis of a set of variables. Variables can be quantitative, qualitative or a mixture of both.

License GPL (>= 2.0)

Depends R (>= 3.0.0)

Imports PCAmixdata

RoxygenNote 7.3.2

Encoding UTF-8

NeedsCompilation no

Repository CRAN

Date/Publication 2025-04-23 18:30:05 UTC

Contents

`bootvar` *Bootstrap of individuals on a numeric matrix and on a categorical matrix*

Description

Draw a bootstrap sample from X.quanti and a bootstrap sample from X.quali

Usage

```
bootvar(X.quanti = NULL, X.quali = NULL)
```

Arguments

`X.quanti`	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).
`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).

clusterscore *Calculates de synthetic variable of a cluster*

Description

Calculates the synthetic variable of a cluster of variables. The variables can be quantitative or qualitative. The synthetic variable is the first principal component of PCAmix. The variance of the synthetic variable is the first eigenvalue. It is equal to the sum of squared correlations or correlation ratios to the synthetic variable. It measures the homogeneity of the cluster.

Usage

```
clusterscore(Z)
```

Arguments

Z	a centered and reduced data matrix obtained with the recod function

Value

f	the synthetic variables i.e. the scores on the first principal component of PCAmix
sv	the standard deviation of f i.e. the first singular value
v	the standardized loadings

Examples

```
data(decathlon)
A <- 1:5
Z <- PCAmixdata::recod(X.quanti=decathlon[1:10,A], X.quali=NULL)$Z
clusterscore(Z)
Z%*%as.matrix(clusterscore(Z)$v)
clusterscore(Z)$f
```

clust_diss *Calculates the aggregation criterion for two clusters of variables*

Description

Calculates the measure of aggregation of two clusters of variables. This measure of aggregation is equal to the decrease in homogeneity for the clusters being merged.

Usage

```
clust_diss(A, B)
```

Arguments

A	a centered and reduced data matrix obtained with the recod function for the first cluster
B	a centered and reduced data matrix obtained with the recod function for the second cluster

Value

The aggregation measure between the two clusters

Examples

```
data(decathlon)
A <- PCAmixdata::recod(X.quanti=decathlon[1:10,1:5], X.quali=NULL)$Z
B <- PCAmixdata::recod(X.quanti=decathlon[1:10,6:10], X.quali=NULL)$Z
clust_diss(A,B)
```

---

clust_diss2 *Dissimilarity between two clusters of variables*

Description

Dissimilarity between two clusters of variables when only the covariance/correlation matrix is known.

Usage

```
clust_diss2(x, A, B)
```

Arguments

x	a covariance/correlation matrix
A	indices of cluster A
B	indices of cluter B

Value

The dissimilarity between the two clusters

Examples

```
data(decathlon)
x <- cor(decathlon[,1:10])
A <- c(1,3,4)
B <- c(2,7,10)
clust_diss2(x,A,B)
```

cutreevar *Cut a tree into groups of variables*

Description

Cuts a hierarchical tree of variables resulting from `hclustvar` into several clusters by specifying the desired number of clusters.

Usage

```
cutreevar(obj, k = NULL, matsim = FALSE)
```

Arguments

`obj`	an object of class 'hclustvar'.
`k`	an integer scalar with the desired number of clusters.
`matsim`	boolean, if TRUE, the matrices of similarities between variables in same cluster are calculated.

Value

`var`	a list of matrices of squared loadings i.e. for each cluster of variables, the squared loadings on first principal component of PCAmix. For quantitative variables (resp. qualitative), squared loadings are the squared correlations (resp. the correlation ratios) with the first PC (the cluster center).
`sim`	a list of matrices of similarities i.e. for each cluster, similarities between their variables. The similarity between two variables is defined as a square cosine: the square of the Pearson correlation when the two variables are quantitative; the correlation ratio when one variable is quantitative and the other one is qualitative; the square of the canonical correlation between two sets of dummy variables, when the two variables are qualitative. `sim` is 'NULL' if 'matsim' is 'FALSE'.
`cluster`	a vector of integers indicating the cluster to which each variable is allocated.
`wss`	the within-cluster sum of squares for each cluster: the sum of the correlation ratio (for qualitative variables) and the squared correlation (for quantitative variables) between the variables and the center of the cluster.
`E`	the pourcentage of homogeneity which is accounted by the partition in k clusters.
`size`	the number of variables in each cluster.
`scores`	a n by k numerical matrix which contains the k cluster centers. The center of a cluster is a synthetic variable: the first principal component calculated by PCAmix. The k columns of `scores` contain the scores of the n observations units on the first PCs of the k clusters.
`coef`	a list of the coefficients of the linear combinations defining the synthetic variable of each cluster.

See Also

hclustvar, summary.clustvar, plot.clustvar, predict.clustvar

Examples

```
data(decathlon)
tree <- hclustvar(decathlon[,1:10])
plot(tree)
#choice of the number of clusters
stability(tree,B=40)
part <- cutreevar(tree,4)
print(part)
summary(part)
```

---

decathlon *Performance in decathlon (data)*

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**.

`getnnsvar` *Nearest neighbor of variables*

Description

Nearest neighbor of variables

Usage

```
getnnsvar(diss, flag)
```

Arguments

`diss`	a dissimilarity matrix between variables
`flag`	a vector of size `p` which indicates if we want to compute nearest neighbor of variable j (flag[j]=1) or not (flag[j]=0)

gironde *gironde*

Description

A list of 4 datasets caracterizing 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 counsil 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]

`hclustvar` *Hierarchical clustering of variables*

Description

Ascendant hierarchical clustering of a set of variables. Variables can be quantitative, qualitative or a mixture of both. The aggregation criterion is the decrease in homogeneity for the clusters being merged. The homogeneity of a cluster is the sum of the correlation ratio (for qualitative variables) and the squared correlation (for quantitative variables) between the variables and the center of the cluster which is the first principal component of PCAmix. PCAmix is defined for a mixture of qualitative and quantitative variables and includes ordinary principal component analysis (PCA) and multiple correspondence analysis (MCA) as special cases. Missing values are replaced by means for quantitative variables and by zeros in the indicator matrix for qualitative variables.

Usage

```
hclustvar(X.quanti = NULL, X.quali = NULL, init = NULL)
```

Arguments

`X.quanti`	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).
`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).
`init`	an initial partition (a vector of integers indicating the cluster to which each variable is allocated).

Details

If the quantitative and qualitative data are in a same dataframe, the function `PCAmixdata::splitmix` can be used to extract automatically the qualitative and the quantitative data in two separated dataframes.

Value

`height`	a set of p-1 non-decreasing real values: the values of the aggregation criterion.
`clusmat`	a p by p matrix with group memberships where each column k corresponds to the elements of the partition in k clusters.
`merge`	a p-1 by 2 matrix. Row i of `merge` describes the merging of clusters at step i of the clustering. If an element j in the row is negative, then observation -j was merged at this stage. If j is positive then the merge was with the cluster formed at the (earlier) stage j of the algorithm. Thus negative entries in `merge` indicate agglomerations of singletons, and positive entries indicate agglomerations of non-singletons.

References

Chavent, M., Liquet, B., Kuentz, V., Saracco, J. (2012), ClustOfVar: An R Package for the Clustering of Variables. Journal of Statistical Software, Vol. 50, pp. 1-16.

See Also

cutreevar, plot.hclustvar, stability

Examples

```
#quantitative variables
data(decathlon)
tree <- hclustvar(X.quanti=decathlon[,1:10], init=NULL)
plot(tree)

#qualitative variables with missing values
data(vnf)
tree_NA <- hclustvar(X.quali=vnf)
plot(tree_NA)
vnf2<-na.omit(vnf)
tree <- hclustvar(X.quali=vnf2)
plot(tree)

#mixture of quantitative and qualitative variables
data(wine)
X.quanti <- PCAmixdata::splitmix(wine)$X.quanti
X.quali <- PCAmixdata::splitmix(wine)$X.quali
tree <- hclustvar(X.quanti,X.quali)
plot(tree)
```

---

hclustvar2 *Hierarchical clustering of variables from a covariance matrix*

Description

Ascendant hierarchical clustering of a set of variables from a covariance/correlation matrix.

Usage

```
hclustvar2(x, init = NULL)
```

Arguments

x	a covariance or correlation matrix.
init	an initial partition (a vector of integers indicating the cluster to which each variable is allocated).

Value

height	a set of p-1 non-decreasing real values: the values of the aggregation criterion.
clusmat	a p by p matrix with group memberships where each column k corresponds to the elements of the partition in k clusters.
merge	a p-1 by 2 matrix. Row i of merge describes the merging of clusters at step i of the clustering. If an element j in the row is negative, then observation -j was merged at this stage. If j is positive then the merge was with the cluster formed at the (earlier) stage j of the algorithm. Thus negative entries in merge indicate agglomerations of singletons, and positive entries indicate agglomerations of non-singletons.

See Also

cutreevar, plot.hclustvar, stability

Examples

```
data(decathlon)
x <- cor(decathlon[,1:10])
tree <- hclustvar2(x)
plot(tree, hang = -1, xlab="", sub="")
```

---

kmeansvar *k-means clustering of variables*

Description

Iterative relocation algorithm of k-means type which performs a partitionning of a set of variables. Variables can be quantitative, qualitative or a mixture of both. The center of a cluster of variables is a synthetic variable but is not a 'mean' as for classical k-means. This synthetic variable is the first principal component calculated by PCAmix. PCAmix is defined for a mixture of qualitative and quantitative variables and includes ordinary principal component analysis (PCA) and multiple correspondence analysis (MCA) as special cases. The homogeneity of a cluster of variables is defined as the sum of the correlation ratio (for qualitative variables) and the squared correlation (for quantitative variables) between the variables and the center of the cluster, which is in all cases a numerical variable. Missing values are replaced by means for quantitative variables and by zeros in the indicator matrix for qualitative variables.

Usage

```
kmeansvar(
  X.quanti = NULL,
  X.quali = NULL,
  init,
  iter.max = 150,
  nstart = 1,
```

```
  matsim = FALSE
)
```

Arguments

X.quanti	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).
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).
init	either the number of clusters or an initial partition (a vector of integers indicating the cluster to which each variable is allocated). If init is a number, a random set of (distinct) columns in X.quali and X.quanti is chosen as the initial cluster centers.
iter.max	the maximum number of iterations allowed.
nstart	if init is a number, nstart corresponds with the number of random sets used in the process.
matsim	boolean, if 'TRUE', the matrices of similarities between variables in same cluster are calculated.

Details

If the quantitative and qualitative data are in a same dataframe, the function splitmix can be used to extract automatically the qualitative and the quantitative data in two separated dataframes.

Value

var	a list of matrices of squared loadings i.e. for each cluster of variables, the squared loadings on first principal component of PCAmix. For quantitative variables (resp. qualitative), squared loadings are the squared correlations (resp. the correlation ratios) with the first PC (the cluster center).
sim	a list of matrices of similarities i.e. for each cluster, similarities between their variables. The similarity between two variables is defined as a square cosine: the square of the Pearson correlation when the two variables are quantitative; the correlation ratio when one variable is quantitative and the other one is qualitative; the square of the canonical correlation between two sets of dummy variables, when the two variables are qualitative. sim is 'NULL if matsim is FALSE.
cluster	a vector of integers indicating the cluster to which each variable is allocated.
wss	the within-cluster sum of squares for each cluster: the sum of the correlation ratio (for qualitative variables) and the squared correlation (for quantitative variables) between the variables and the center of the cluster.
E	the pourcentage of homogeneity which is accounted by the partition in k clusters.
size	the number of variables in each cluster.
scores	a n by k numerical matrix which contains the k cluster centers. The center of a cluster is a synthetic variable: the first principal component calculated by PCAmix. The k columns of scores contain the scores of the n observations units on the first PCs of the k clusters.

coef	a list of the coefficients of the linear combinations defining the synthetic variable of each cluster.

References

Chavent, M., Liquet, B., Kuentz, V., Saracco, J. (2012), ClustOfVar: An R Package for the Clustering of Variables. Journal of Statistical Software, Vol. 50, pp. 1-16.

See Also

summary.clustvar, plot.clustvar, predict.clustvar

Examples

```
data(decathlon)
#choice of the number of clusters
tree <- hclustvar(X.quanti=decathlon[,1:10])
stab <- stability(tree,B=60)
#a random set of variables is chosen as the initial cluster centers, nstart=10 times
part1 <- kmeansvar(X.quanti=decathlon[,1:10],init=5,nstart=10)
summary(part1)
#the partition from the hierarchical clustering is chosen as initial partition
part_init<-cutreevar(tree,5)$cluster
part2<-kmeansvar(X.quanti=decathlon[,1:10],init=part_init,matsim=TRUE)
summary(part2)
part2$sim
```

---

mixedVarSim *Similarity between two variables*

Description

Returns the similarity between two quantitative variables, two qualitative variables or a quantitative variable and a qualitative variable. The similarity between two variables is defined as a square cosine: the square of the Pearson correlation when the two variables are quantitative; the correlation ratio when one variable is quantitative and the other one is qualitative; the square of the canonical correlation between two sets of dummy variables, when the two variables are qualitative.

Usage

```
mixedVarSim(X1, X2)
```

Arguments

X1	a vector or a factor
X2	a vector or a factor

plot.clustab *Plot of an index of stability of partitions of variables*

Description

Plot of the index of stability of the partitions against the number of clusters.

Usage

```
## S3 method for class 'clustab'
plot(x, nmin = NULL, nmax = NULL, ...)
```

Arguments

x	an object of class clustab.
nmin	the minimum number of clusters in the plot.
nmax	the maximum number of clusters in the plot.
...	further arguments passed to or from other methods.

See Also

stability

Examples

```
data(decathlon)
tree <- hclustvar(X.quanti=decathlon[,1:10])
stab<-stability(tree,B=20)
plot(stab,nmax=7)
```

plot.clustvar *Plot loadings in each cluster.*

Description

Plot dotchart with the "loadings" of the variables in each cluster. The loading of a numerical variable is the correlation between this variables and the synthetic variable of its cluster. The loading of the level of a categorical variable is the mean value of the synthetic variable of the cluster on observations having this level.

Usage

```
## S3 method for class 'clustvar'
plot(x, ...)
```

Arguments

x	an object of class clustvar obtained with cutreevar or kmeansvar.
...	Further arguments to be passed to or from other methods. They are ignored in this function.

Value

coord.quanti	coordinates of quantitative variables belonging to cluster k on the synthetic variable associate to the same cluster k
coord.levels	coordinates of levels of categorical variables belonging to cluster k on the synthetic variable associate to the same cluster k

Examples

```
data(wine)
X.quanti <- PCAmixdata::splitmix(wine)$X.quanti
X.quali <- PCAmixdata::splitmix(wine)$X.quali
tree <- hclustvar(X.quanti,X.quali)
tree.cut<-cutreevar(tree,6)

#plot of scores on synthetic variables
res.plot <- plot(tree.cut)
res.plot$coord.quanti
res.plot$coord.levels
```

---

plot.hclustvar *Dendrogram of the hierarchy of variables*

Description

Dendrogram of the hierarchy of variables resulting from hclustvar and aggregation levels plot.

Usage

```
## S3 method for class 'hclustvar'
plot(x, type = "tree", sub = "", ...)
```

Arguments

x	an object of class hclustvar.
type	if type="tree" plot of the dendrogram and if type="index" aggregation levels plot.
sub	a sub title for the plot.
...	further arguments passed to or from other methods.

See Also

hclustvar

Examples

```
data(wine)
X.quanti <- PCAmixdata::splitmix(wine)$X.quanti
X.quali <- PCAmixdata::splitmix(wine)$X.quali
tree <- hclustvar(X.quanti,X.quali)
plot(tree)

#Aggregation levels plot
plot(tree,type="index")
```

---

predict.clustvar *Scores of new objects on the synthetic variables of a given partition*

Description

A partition of variables obtained with kmeansvar or with cutreevar is given in input. Each cluster of this partition is associated with a synthetic variable which is a linear combination of the variables of the cluster. The coefficients of these k linear combinations (one for each cluster) are used here to calculate new scores of a objects described in a new dataset (with the same variables). The output is the matrix of the scores of these new objects on the k synthetic variables.

Usage

```
## S3 method for class 'clustvar'
predict(object, X.quanti = NULL, X.quali = NULL, ...)
```

Arguments

object	an object of class clustvar
X.quanti	numeric matrix of data for the new objects
X.quali	a categorical matrix of data for the new objects
...	Further arguments to be passed to or from other methods. They are ignored in this function.

Value

Returns the matrix of the scores of the new objects on the k syntetic variables of the k-clusters partition given in input.

Examples

```
data(wine)
n <- nrow(wine)
sub <- 10:20
data.sub <- wine[sub,] #learning sample
X.quanti <- wine[sub,c(3:29)] #learning sample
X.quali <- wine[sub,c(1,2)]
part <-kmeansvar(X.quanti, X.quali, init=5)
X.quanti.t <- wine[-sub,c(3:29)]
X.quali.t <- wine[-sub,c(1,2)]
new <- predict(part,X.quanti.t,X.quali.t)
```

---

protein *Protein data*

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).

---

rand *Rand index between two partitions*

Description

Returns the Rand index, the corrected Rand index or the asymmetrical Rand index. The asymmetrical Rand index (corrected or not) measures the inclusion of a partition P into and partition Q with the number of clusters in P greater than the number of clusters in Q.

Usage

```
rand(P, Q, symmetric = TRUE, adj = TRUE)
```

Arguments

P	a factor, e.g., the first partition.
Q	a factor, e.g., the second partition.
symmetric	a boolean. If FALSE the asymmetrical Rand index is calculated.
adj	a boolean. If TRUE the corrected index is calculated.

See Also

stability

---

selvar *Selection of a given number of variables in each cluster.*

Description

This function selects in each cluster a given number of variables having the highest squared loadings. The squared loading of a variable in a cluster is its squared correlation (for numerical variable) and its correlation ratio (for categorical variable) with the first PC of PCAmix applied to the variables of the cluster.

Usage

```
selvar(part, nsel)
```

Arguments

part	an object of class clustvar
nsel	the number of variables selected in each cluster.

Details

If the number of variables in a cluster is smaller than nsel, all the variables of the cluster are selected

Value

Returns a list where each element contains the nsel selected variables.

Examples

```
data(decathlon)
tree <- hclustvar(decathlon[,1:10])
part <- cutreevar(tree,4)
part$var
selvar(part,2)
```

stability *Stability of partitions from a hierarchy of variables*

Description

Evaluates the stability of partitions obtained from a hierarchy of p variables. This hierarchy is performed with hclustvar and the stability of the partitions of 2 to p-1 clusters is evaluated with a bootstrap approach. The boostrap approch is the following: hclustvar is applied to B boostrap samples of the n rows. The partitions of 2 to p-1 clusters obtained from the B bootstrap hierarchies are compared with the partitions from the initial hierarchy . The mean of the corrected Rand indices is plotted according to the number of clusters. This graphical representation helps in the determination of a suitable numbers of clusters.

Usage

```
stability(tree, B = 100, graph = TRUE)
```

Arguments

tree	an object of class hclustvar.
B	the number of bootstrap samples.
graph	boolean, if 'TRUE' a graph is displayed.

Value

matCR	matrix of corrected Rand indices.
meanCR	vector of mean corrected Rand indices.

See Also

plot.clustab, hclustvar

Examples

```
data(decathlon)
tree <- hclustvar(X.quanti=decathlon[,1:10])
stab<-stability(tree,B=20)
plot(stab,nmax=7)
boxplot(stab$matCR[,1:7])
```

summary.clustab *Summary of a 'clustab' object*

Description

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

Usage

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

Arguments

`object`	An object of class `clustab` generated by the function `stability`.
`...`	Further arguments passed to or from other methods.

See Also

`stability`

summary.clustvar *Summary of a 'hclustvar' object*

Description

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

Usage

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

Arguments

`object`	an object of class `clustvar`.
`...`	further arguments passed to or from other methods.

Value

Returns a list of matrices of squared loadings i.e. for each cluster of variables, the squared loadings on first principal component of PCAmix. For quantitative variables (resp. qualitative), squared loadings are the squared correlations (resp. the correlation ratios) with the first PC (the cluster center). If the partition of variables has been obtained with `kmeansvar` the number of iteration until convergence is also indicated.

See Also

kmeansvar, cutreevar

Examples

```
data(decathlon)
part<-kmeansvar(X.quanti=decathlon[,1:10],init=5)
summary(part)
```

---

vnf *User satisfaction survey with 1232 individuals and 14 questions*

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 *Wine*

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