Package 'CAvariants'

January 20, 2025

Type Package

Title Correspondence Analysis Variants

Version 6.0

Date 2023-10-19

Author Rosaria Lombardo and Eric J Beh

Maintainer Rosaria Lombardo <rosaria.lombardo@unicampania.it>

Description Provides six variants of two-way correspondence analysis (ca): simple ca, singly ordered ca, doubly ordered ca, non symmetrical ca, singly ordered non symmetrical ca, and doubly ordered non symmetrical ca.

Depends R (> 3.0.1), methods, tools, ggforce, ggrepel, gridExtra

Imports ggplot2, plotly

LazyData true

License GPL (> 2)

URL https://www.R-project.org

NeedsCompilation no

Repository CRAN

Date/Publication 2023-10-19 17:20:06 UTC

Contents

`asbestos` *Selikoff's data, a two-way contingency table.*

Description

The data set consists of 4 rows and 5 columns. The rows represent the degree of severity of asbestosis and the columns are concerned with the time of exposure to asbestos in years of the workers

Usage

```
data(asbestos)
```

Format

The format is:
row names [1:4] "None" "grade1" "grade2" "grade3"
col names [1:5] "0-9" "10-19" "20-29" "30-39" "40+"

References

Beh EJ and Lombardo R 2014 Correspondence Analysis: Theory, Practice and New Strategies. Wiley.
Selikoff IJ 1981 Household risks with inorganic fibers. Bulletin of the New York Academy of Medicine, 57, 947 – 961.

Examples

```
asbestos <-structure(c(310, 36, 0, 0, 212, 158, 9, 0, 21, 35, 17, 4, 25,
102, 49, 18, 7, 35, 51, 28), .Dim = 4:5, .Dimnames = list(c("none",
"grade1", "grade2", "grade3"), c("0-9", "10-19", "20-29", "30-39",
"40+")))
dim(asbestos)
dimnames(asbestos)
```

cabasic *Classical two-way correspondence analysis*

Description

This function is used in the main function CAvariants when the input parameter is catype = "CA". It performs the singular value decomposition of Pearson's ratio and computes principal axes, coordinates, the weights of rows and columns, the total inertia (equal to Pearson's index) and the rank of the matrix.

Usage

```
cabasic(Xtable)
```

Arguments

Xtable	The two-way contingency table.

Note

This function belongs to the R object class called cabasicresults.

Author(s)

Rosaria Lombardo and Eric J. Beh

References

Beh EJ and Lombardo R 2014 Correspondence Analysis: Theory, Practice and New Strategies. Wiley.
Lombardo R Beh EJ 2016 Variants of Simple Correspondence Analysis. The R Journal, 8 (2), 167–184.
Lombardo R Beh EJ and Kroonenberg PM 2016 Modelling Trends in Ordered Correspondence Analysis Using Orthogonal Polynomials. Psychometrika, 81(2), 325–349.

Examples

```
data(asbestos)
cabasic(asbestos)
```

caplot3d *Three dimensional correspondence plot*

Description

This function is used in the plot function `plot.CAvariants` when the logical parameter is `plot3d = TRUE`. It produces a 3-dimensional visualization of the association.

Usage

```
caplot3d(coordR, coordC, inertiaper, firstaxis = 1, lastaxis = 2, thirdaxis = 3)
```

Arguments

`coordR`	The row principal or standard coordinates.
`coordC`	The column principal or standard coordinates.
`inertiaper`	The percentage of the total inertia explained inertia by each dimension.
`firstaxis`	The first axis number. By default, `firstaxis = 1`.
`lastaxis`	The second axis number. By default, `lastaxis = 2`.
`thirdaxis`	The third axis number. By default, `thirdaxis = 3`.

Note

This function depends on the `R` library `plotly`.

Author(s)

Rosaria Lombardo and Eric J. Beh

References

Beh EJ and Lombardo R 2014 Correspondence Analysis: Theory, Practice and New Strategies. Wiley.
Lombardo R Beh EJ 2016 Variants of Simple Correspondence Analysis. The R Journal, 8 (2), 167–184.

`caplotord` *Row isometric or column isometric biplot for ordered variants of correspondence analysis*

Description

This function is used in the main plot function when the plot type parameter is `plottype = "biplot"`. It can produce a row polynomial biplot or a column polynomial biplot.

Usage

```
caplotord(frows, gcols, firstaxis, lastaxis, nseg, inertiapc, thingseg, col1,
col2, col3, size1, size2)
```

Arguments

`frows`	The row principal or standard coordinates.
`gcols`	The column principal or standard coordinates.
`firstaxis`	The first polynomial axis number.
`lastaxis`	The second polynomial axis number.
`nseg`	The vectors/arrows number where to project principal (or standard) coordinates.
`inertiapc`	The percentage of the explained inertia by each dimension.
`thingseg`	The principal or standard coordinates used to draw vectors (arrows).
`col1`	The colour for the row variable labels.
`col2`	The colour for the column variable labels.
`col3`	The colour for the vectors (arrows) used in biplots.
`size1`	The size of the plotted symbol for categories in biplot.
`size2`	The size of the plotted text for categories in biplot.

Note

This function depends on the `R` library `plotly`.

Author(s)

Rosaria Lombardo and Eric J. Beh

References

Beh EJ and Lombardo R 2014 Correspondence Analysis: Theory, Practice and New Strategies. Wiley.
Lombardo R Beh EJ 2016 Variants of Simple Correspondence Analysis. The R Journal, 8 (2), 167–184.

`caRbiplot` *Row isometric biplot or Column isometric biplot*

Description

This function is used in the main plot function when the plot type parameter is `plottype = "biplot"`. It can produce a row biplot or a column biplot.

Usage

```
caRbiplot(frows, gcols, firstaxis, lastaxis, inertiapc,  bip="row", size1,size2)
```

Arguments

`frows`	The row principal or standard coordinates.
`gcols`	The column principal or standard coordinates.
`firstaxis`	The first axis number.
`lastaxis`	The second axis number.
`inertiapc`	The percentage of the explained inertia by each dimension.
`bip`	The type of biplot. One may specify a row-isometric biplot or a column-isometric biplot (when using in the function `plot.CAvariant` the parameter `biptype = "row"` or `biptype = "column"`).
`size1`	The size of the plotted symbol for categories in biplot.
`size2`	The size of the plotted text for categories in biplot.

Author(s)

Rosaria Lombardo and Eric J. Beh

References

Beh EJ and Lombardo R 2014 Correspondence Analysis: Theory, Practice and New Strategies. Wiley.
Lombardo R Beh EJ 2016 Variants of Simple Correspondence Analysis. The R Journal, 8 (2), 167–184.

`CAvariants` *Six variants of correspondence analysis*

Description

It performs
1) simple correspondence analysis
2) doubly ordered correspondence analysis
3) singly ordered correspondence analysis
4) non symmetrical correspondence analysis
5) doubly ordered non symmetrical correspondence analysis
6) singly ordered non symmetrical correspondence analysis

Usage

```
CAvariants(Xtable, mj = NULL, mi = NULL, firstaxis = 1, lastaxis = 2,
catype = "CA",  M = min(nrow(Xtable), ncol(Xtable)) - 1, alpha = 0.05)
```

Arguments

`Xtable`	The two-way contingency table.
`mi`	The assigned ordered scores for the row categories. By default, `mi = NULL`, which gives consecutive integer valued (natural) scores.
`mj`	The assigned ordered scores for the column categories, By default, `mj = NULL`, which gives consecutive integer valued (natural) scores.
`firstaxis`	The horizontal polynomial, or principal, axis. It is used for the construction of the Inner product table. By default `firstaxis = 1`.
`lastaxis`	The vertical polynomial, or principal, axis. It is used for the construction of the Inner product table. By default `lastaxis = 2`.
`catype`	The input parameter for specifying what variant of correspondence analysis is to be performed. By default, `catype = "CA"`. Other possible values are: `catype = "SOCA"`, `catype = "DOCA"`, `catype = "NSCA"`, `catype = "SONSCA"`, `catype = "DONSCA"`.
`M`	The number of axes used for determining the structure of the elliptical confidence regions. By default, `M = min(nrow(Xtable), ncol(Xtable)) - 1`, i.e. the rank of the data matrix.
`alpha`	The level of significance for the elliptical regions. By default, `alpha = 0.05`.

Value

Description of the output returned

`Xtable`	The two-way contingency table.
`rows`	The number of rows of the two-way contingency table.
`cols`	The number of columns of the two-way contingency table.

r	The rank of the two-way contingency table.
n	The total number of observations of the two-way contingency table.
rowlabels	The labels of the row variable.
collabels	The labels of the column variable.
Rprinccoord	The row principal coordinates. When the input parameter catype is "DOCA", "SOCA", "DONSCA" or "SONSCA", they are row principal polynomial coordinates.
Cprinccoord	The column principal coordinates. When the input parameter catype is "DOCA", "SOCA", "DONSCA" or "SONSCA", they are column principal polynomial coordinates.
Rstdcoord	The row standard coordinates. When the input parameter catype is "DOCA", "SOCA", "DONSCA" or "SONSCA", they are row standard polynomial coordinates.
Cstdcoord	The column standard coordinates. When the input parameter catype is "DOCA", "SOCA", "DONSCA" or "SONSCA", they are column standard polynomial coordinates.
tauden	The denominator of the Goodman-Kruskal tau index is given when the input parameter catype is "NSCA", "SONSCA", or "DONSCA". Otherwise it is NULL.
tau	The index of Goodman and Kruskal is given when the input parameter catype is "NSCA", "SONSCA", or "DONSCA". Otherwise it is NULL.
inertiasum	The total inertia of the analysis based on Pearson's chi-squared when catype is "CA", "DOCA" or "SOCA", or based on the Goodman-Kruskal tau when catype is "NSCA", "DONSCA" or "SONSCA" (numerator of the Goodman-Kruskal tau index).
singvalue	The singular values of the two-way contingency table.
inertias	The inertia in absolute value and percentage, in the row space for each principal or polynomial axis.
inertias2	The inertia in absolute value and percentage, in the column space for each principal or polynomial axis. When catype is "CA" or "NSCA" the associated inertia in the row and column spaces are the same for each principal axis.
t.inertia	The total inertia of the two-way contingency table.
comps	The polynomial components of inertia when the variables are ordered.
catype	The type of correspondence analysis chosen by the analyst. By default, catype = "ca".
mj	The ordered scores of the column variable. When mj = NULL, the natural scores are used (i = 1,...,cols).
mi	The ordered scores of the row variable. When mi = NULL, the natural scores are used (i = 1,...,rows).
pcc	The weighted centered column profile matrix.
Jmass	The weight matrix of the column variable.
Imass	The weight matrix of the row variable.
Innprod	The inner product, Inner product, of the biplot coordinates (for the two axes defined by firstaxis = 1 and lastaxis = 2)

Z	The generalised correlation matrix when catype = "SOCA", catype = "DOCA" , catype = "SONSCA", catype = "DONSCA", but when catype = "CA", or catype = "NSCA", it gives again the inner product matrix of biplot coordinates.
M	The number of axes used for determining the structure of the elliptical confidence regions. By default, M = min(nrow(Xtable), ncol(Xtable)) - 1, i.e. the rank of the data matrix.
eccentricity	When ellcomp = TRUE, the output gives the eccentricity of the confidence ellipses.
row.summ	When ellcomp = TRUE, the output gives for each row the summary results that contain the semi-major axis length of the ellipse, HL Axis 1, the semi-minor axis length for the ellipse, HL Axis 2, the area of the ellipse, Area and the p-value, P-value.
col.summ	When ellcomp = TRUE, the output gives for each column point, the summary results that contain the semi-major axis length of the ellipse, HL Axis 1, the semi-minor axis length for the ellipse, HL Axis 2, the area of the ellipse, Area and the p-value, P-value.

Note

This function recalls internally many other functions, depending on the setting of the input parameter catype, it recalls one of the six functions which does a variant of correspondence analysis. After performing a variant of correspondence analysis, it gives the output object necessary for printing and plotting the results. These two important functions are print.CAvariants and plot.CAvariants.

Author(s)

Rosaria Lombardo and Eric J Beh

References

Beh EJ and Lombardo R 2014 Correspondence Analysis: Theory, Practice and New Strategies. Wiley.

Lombardo R Beh EJ 2016 Variants of Simple Correspondence Analysis. The R Journal, 8 (2), 167–184.

Lombardo R Beh EJ and Kroonenberg PM 2016 Modelling Trends in Ordered Correspondence Analysis Using Orthogonal Polynomials. Psychometrika, 81(2), 325–349.

Examples

```
data(asbestos)
CAvariants(asbestos, catype = "CA")
CAvariants(asbestos, catype = "DOCA", mi = c(1:nrow(asbestos)), mj =c(4.5,14.5,24.5,34.5,44.5),
firstaxis = 1, lastaxis = 2, M = min(nrow(asbestos), ncol(asbestos)) - 1)
CAvariants(asbestos, catype = "DONSCA")
data(shopdataM)
CAvariants(shopdataM, catype = "NSCA")
CAvariants(shopdataM, catype = "SONSCA")
CAvariants(shopdataM, catype = "SOCA")
```

`compsonetable.exe` *Polynomial component of inertia in column space*

Description

This function allows the analyst to compute the contribution that the polynomial components make to the inertia (Pearson's chi-squared statistic or the Goodman-Kruskal tau index). The ordered variable should be the column variable that is transformed by polynomials. The polynomial components are the column polynomial components. The given input matrix is the Z matrix of generalised correlations from the hybrid decomposition. It is called by `CAvariants` when `catype = "SOCA"` or `catype = "SONSCA"`.

Usage

```
compsonetable.exe(Z)
```

Arguments

`Z`	The matrix of generalised correlations between the polynomial and principal axes.

Value

The value returned is the matrix

`comps`	The matrix of the column polynomial component of inertia.

Note

This function belongs to the class called `cacorporateplus`.

Author(s)

Rosaria Lombardo and Eric J. Beh

References

Beh EJ and Lombardo R 2014 Correspondence Analysis: Theory, Practice and New Strategies. Wiley.

Lombardo R Beh EJ 2016 Variants of Simple Correspondence Analysis. The R Journal, 8 (2), 167–184.

Lombardo R Beh EJ and Kroonenberg PM 2016 Modelling Trends in Ordered Correspondence Analysis Using Orthogonal Polynomials. Psychometrika, 81(2), 325–349.

compstable.exe *Polynomial component of inertia for the row and column spaces*

Description

This function allows the analyst to compute the contribution of the polynomial components to the inertia (Pearson's chi-squared statistic or the Goodman-Kruskal tau index). The ordered variable should be both the row and column variables that are transformed by the polynomials. The polynomial components are the row and column polynomial components. The given input matrix is the Z matrix of generalised correlations from the bivariate moment decomposition. It is called by `CAvariants` when `catype = "DOCA"` or `catype = "DONSCA"`.

Usage

```
compstable.exe(Z)
```

Arguments

`Z`	The matrix of generalised correlations between the polynomial axes.

Value

The value returned is the matrix

`comps`	The matrix of the polynomial components of the inertia.

Author(s)

Rosaria Lombardo and Eric J. Beh

References

Beh EJ and Lombardo R 2014 Correspondence Analysis: Theory, Practice and New Strategies. Wiley.

Lombardo R Beh EJ 2016 Variants of Simple Correspondence Analysis. The R Journal, 8 (2), 167–184.

Lombardo R Beh EJ and Kroonenberg PM 2016 Modelling Trends in Ordered Correspondence Analysis Using Orthogonal Polynomials. Psychometrika, 81(2), 325–349.

`docabasic` *Doubly, or two-way, ordered correspondence analysis: for two ordered variables*

Description

This function is used by the main function `CAvariants` when the input parameter is `catype = "DOCA"`. It performs the bivariate moment decomposition of the Pearson ratio, computes polynomial axes, coordinates, weights of rows and columns, total inertia (based on Pearson's chi-squared statistic), the rank of the matrix. It also decomposes the inertia into row and column polynomial components.

Usage

```
docabasic(Xtable, mi, mj)
```

Arguments

`Xtable`	The two-way contingency table.
`mi`	The set of ordered row scores. By default, `mi = c(1:nrow(Xtable))` (natural scores).
`mj`	The set of ordered column scores. By default, `mj = c(1:ncol(Xtable))` (natural scores).

Author(s)

Rosaria Lombardo and Eric J. Beh

References

Beh EJ and Lombardo R 2014 Correspondence Analysis: Theory, Practice and New Strategies. Wiley.

Lombardo R Beh EJ 2016 Variants of Simple Correspondence Analysis. The R Journal, 8 (2), 167–184.

Lombardo R Beh EJ and Kroonenberg PM 2016 Modelling Trends in Ordered Correspondence Analysis Using Orthogonal Polynomials. Psychometrika, 81(2), 325–349.

Examples

```
data(asbestos)
mi <- c(1,2,3,4) #natural scores for rows
mj <- c(4.5,14.5,24.5,34.5,44.5) #midpoints for columns
docabasic(asbestos, mi, mj)
```

`donscabasic` *Doubly, or two-way ordered, non symmetrical correspondence analysis: for two ordered variables*

Description

This function is used in the main function `CAvariants` when the input parameter is `catype = "DONSCA"`. It performs the bivariate moment decomposition of the numerator of the Goodman-Kruskal tau index for a contingency table consisting of two ordered variables. It computes the polynomial axes, coordinates, weights of the rows and columns, total inertia (equal to the numerator of the tau index) and the rank of the matrix. It also decomposes the inertia into row and column polynomial components.

Usage

```
donscabasic(Xtable, mi, mj)
```

Arguments

`Xtable`	The two-way contingency table.
`mi`	The set of ordered row scores. By default, `mi = c(1:nrow(Xtable))` (natural scores).
`mj`	The set of ordered column scores. By default, `mj = c(1:ncol(Xtable))` (natural scores).

Author(s)

Rosaria Lombardo and Eric J. Beh

References

Beh EJ and Lombardo R 2014 Correspondence Analysis: Theory, Practice and New Strategies. Wiley.
Lombardo R Beh EJ 2016 Variants of Simple Correspondence Analysis. The R Journal, 8 (2), 167–184.
Lombardo R Beh EJ and Kroonenberg PM 2016 Modelling Trends in Ordered Correspondence Analysis Using Orthogonal Polynomials. Psychometrika, 81(2), 325–349.

Examples

```
data(asbestos)
mi <- c(1, 2, 3, 4) # natural scores for the rows
mj <- c(4.5, 14.5, 24.5, 34.5, 44.5) #midpoints for the columns
donscabasic(asbestos, mi, mj)
```

`emerson.poly` *Orthogonal polynomials*

Description

This function is called from the functions `docabasic, socabasic, sonscabasic` and `donscabasic`. It computes the orthogonal polynomials for the ordered categorical variables. The number of the polynomials is equal to the number of categories for that variable less one. The function computes the polynomial transformation of the ordered categorical variable.

Usage

```
emerson.poly(mj, pj)
```

Arguments

`mj`	The ordered scores of the ordered variable. By default, `mj = NULL`, the natural scores (1, 2, ...) are computed.
`pj`	The marginal relative frequencies of the ordered variable.

Value

Describe the value returned

`B`	the matrix of the orthogonal polynomials with the trivial polynomial removed.

Note

Note that the sum of the marginal relative frequencies of the ordered variables must be one.

Author(s)

Rosaria Lombardo and Eric J Beh

References

Beh EJ and Lombardo R 2014 Correspondence Analysis: Theory, Practice and New Strategies. Wiley.
Emerson PL 1968 Numerical construction of orthogonal polynomials from a general recurrence formula. Biometrics, 24 (3), 695-701.
Lombardo R Beh EJ 2016 Variants of Simple Correspondence Analysis. The R Journal, 8 (2), 167–184.
Lombardo R Beh EJ and Kroonenberg PM 2016 Modelling Trends in Ordered Correspondence Analysis Using Orthogonal Polynomials. Psychometrika, 81(2), 325-349.

nscabasic *Two-way non symmetrical correspondence analysis*

Description

This function is used in the main function `CAvariants` when the input parameter is `catype = "NSCA"`. It calculates the singular value decomposition of the numerator of the Goodman-Kruskal tau index (index of predictability), computes principal axes, coordinates, weights of the rows and columns, total inertia (numerator of the tau index) and the rank of the matrix.

Usage

```
nscabasic(Xtable)
```

Arguments

`Xtable`	The two-way contingency table.

Author(s)

Rosaria Lombardo and Eric J. Beh

References

Beh EJ and Lombardo R 2014 Correspondence Analysis: Theory, Practice and New Strategies. Wiley.
Lombardo R Beh EJ 2016 Variants of Simple Correspondence Analysis. The R Journal, 8 (2), 167–184.
Lombardo R Beh EJ and Kroonenberg PM 2016 Modelling Trends in Ordered Correspondence Analysis Using Orthogonal Polynomials. Psychometrika, 81(2), 325–349.

Examples

```
data(asbestos)
nscabasic(asbestos)
```

plot.CAvariants *Main plot function*

Description

This function produces the graphical display for the selected variant of correspondence analysis. When catype = "CA" catype = "NSCA" and plottype = "classic", the function produces a plot of the principal coordinates for the row and column categories.
When plottype = "biplot", it produces a biplot graphical display, or a polynomial biplot in case of ordered variables. For an ordered analysis only the polynomial biplots are constructed. In particular, for the singly ordered variants only the row isometric polynomial biplot is appropriate. When the parameter catype defines an ordered variant of CA, the input parameter plottype should be equal to plottype = "biplot". If biptype = "row", it will produce a row isometric polynomial biplot.

Usage

```
## S3 method for class 'CAvariants'
plot(x, firstaxis = 1, lastaxis = 2, thirdaxis = 3, cex = 0.8,
cex.lab = 0.8, plottype = "biplot", biptype = "row",
scaleplot = 1, posleg = "right", pos = 2, ell = FALSE,
alpha = 0.05, plot3d = FALSE, size1 = 1.5, size2 = 3, ...)
```

Arguments

x	The name of the output object used with the main function CAvariants.
firstaxis	The horizontal polynomial, or principal, axis. By default, firstaxis = 1.
lastaxis	The vertical polynomial, or principal, axis. By default, lastaxis = 2.
thirdaxis	The third polynomial, or principal, axis in tridimensional plot. By default, thirdaxis = 3.
cex	The parameter for setting the size of the character labels for the points in a graphical display. By default, cex = 0.8.
cex.lab	The parameter for setting the size of the character labels of axes in graphical displays. By default, cex.lab = 0.8.
plottype	The type of graphical display required (either a correspondence plot or a biplot). The type of graphical display to be constructed. By default, plottype = "biplot"; the alternative is plottype = "classic".
biptype	The parameter for specifying the type of biplot. One may specify a row-isometric biplot (biptype = "row") or a column-isometric biplot (biptype = "column"). This feature is available for the nominal symmetrical and non-symmetrical correspondence analyses. By default, a row-isometric biplot, biptype = "row", is produced.
scaleplot	The parameter for scaling the classic plot and biplot coordinates. See Gower et al. (2011), section 2.3.1, or page 135 of Beh and Lombardo (2014). By default, scaleplot = 1.
posleg	The position of the legend when portraying trends of the categories for ordered variants of correspondence analysis. By default, posleg = "right".
pos	The parameter that specifies the position of label of each point in the graphical display. By default, pos = 2.

ell	The logical parameter which specifies whether algebraic confidence ellipses are to be included in the plot or not. Setting the input parameter to ell = TRUE will assess the statistical significance of each category to the association between the variables. By default, ell = FALSE.
alpha	The confidence level of the elliptical regions. By default, alpha = 0.05.
plot3d	The logical parameter specifies whether a 3D plot is to be included in the output or not. By default, plot3d = FALSE.
size1	The size of the plotted symbol. By default, size = 1.5.
size2	The size of the plotted text. By default, size = 3.
...	Further arguments passed to, or from, other functions.

Details

It produces either a classical or biplot graphical display. Further, when catype = "DOCA", catype = "SOCA", catype = "DONSCA" or catype = "SONSCA", the trends of the row and column variables (after the reconstruction of column profiles by the polynomials) is portrayed.
For classical biplot displays, it superimposes the algebraic confidence ellipses. It uses the secondary plot function caellipse (or nscaellipse) for the symmetrical (or non symmetrical) CA variants.

Note

For the classical plots, row and column principal coordinates are plotted. For biplots, one set of coordinates is the standard coordinates and the other is the principal coordinates. When an ordered variant of correspondence analysis is performed, the biplot is constructed where one set of coordinates consists of the standard polynomial coordinates and the other one is the principal polynomial coordinates.

Author(s)

Rosaria Lombardo and Eric J Beh

References

Beh EJ and Lombardo R 2014 Correspondence Analysis: Theory, Practice and New Strategies. Wiley.
Gower J, Lubbe S, and le Roux, N 2011 Understanding Biplots. Wiley.
Lombardo R Beh EJ 2016 Variants of Simple Correspondence Analysis. The R Journal, 8 (2), 167–184.
Lombardo R Beh EJ and Kroonenberg PM 2016 Modelling Trends in Ordered Correspondence Analysis Using Orthogonal Polynomials. Psychometrika, 81(2), 325–349.

Examples

```
data(asbestos)
resasbestosCA<-CAvariants(asbestos, catype = "CA", M=2)
plot(resasbestosCA, plottype = "classic", plot3d = TRUE)
plot(resasbestosCA, plottype = "classic",  ell = TRUE)
plot(resasbestosCA, plottype = "biplot", biptype = "column", scaleplot=1.5)
resasbestosDOCA<-CAvariants(asbestos, catype = "DOCA")
```

```
plot(resasbestosDOCA, plottype = "biplot", biptype = "column")
resasbestosNSCA<-CAvariants(asbestos, catype = "NSCA")
plot(resasbestosNSCA, plottype = "biplot", biptype = "column", plot3d = TRUE)
```

`print.CAvariants` *Main printing function for numerical summaries*

Description

This function prints the numerical output for any of the six variants of correspondence analysis called by `catype`.
The input parameter is the name of the output of the main function `CAvariants`.

Usage

```
## S3 method for class 'CAvariants'
print(x, printdims = 2, ellcomp = TRUE, digits = 3,...)
```

Arguments

`x`	The name of the output object from the main function `CAvariants`.
`printdims`	The number of dimensions that are used for summarising the numerical output of the analysis. By default, `printdims = 2`. the maximum number is equal to the rank of the table.
`ellcomp`	This parameter specifies whether the characteristics of the confidence ellipses (eccentricity, semi-axis, area, p-values) are to be computed. By default, `ellcomp = TRUE`.
`digits`	The number of decimal places used for displaying the numerical summaries of the analysis. By default, `digits = 3`.
`...`	Further arguments passed to, or from, other functions.

Details

This function uses another function (called `printwithaxes`) for specifying the number of columns of a matrix to print.

Value

The output returned depends on the type of correspondence analysis that is performed

`Xtable`	The two-way contingency table.
`Row weights: Imass`	The row weight matrix. These weights depend on the type of analysis that is performed.
`Column weights: Jmass`	The column weight matrix. These weights are equal to the column marginal relative frequencies for all types of analysis performed.

`Total inertia` The total inertia of the analysis performed. For example, for variants of non symmetrical correspondence analysis, the output produced includes the numerator of the Goodman-Kruskal tau index, its C-statistic and p-value.

`Inertias` The inertia values, their percentage contribution to the total inertia and the cumulative percent inertias for the row and column variables.

`Generalised correlation matrix`
The matrix of generalised correlations when performing an ordered correspondence analysis, `catype` must be `"DOCA"`, `"DONSCA"`, `"SOCA"` or `"SONSCA"`.

`Row principal coordinates`
The row principal coordinates when `catype = "CA"` or `catype = "NSCA"`.

`Column principal coordinates`
The column principal coordinates when `catype = "CA"` or `catype = "NSCA"`.

`Row standard coordinates`
The row standard coordinates when `catype = "CA"` or `catype = "NSCA"`.

`Column standard coordinates`
The column standard coordinates when `catype = "CA"` or `catype = "NSCA"`.

`Row principal polynomial coordinates`
The row principal polynomial coordinates when performing an ordered correspondence analysis.

`Column principal polynomial coordinates`
The column principal coordinates when performing a doubly ordered correspondence analysis.

`Row standard polynomial coordinates`
The row standard polynomial coordinates, when performing an ordered variant of correspondence analysis.

`Column standard polynomial coordinates`
The column standard polynomial coordinates, when performing an ordered variant of correspondence analysis.

`Row distances from the origin of the plot`
The squared Euclidean distance of the row categories from the origin of the plot.

`Column distances from the origin of the plot`
The squared Euclidean distance of the column categories from the origin of the plot.

`Polynomial components`
The polynomial components of the total inertia and their p-values. The total inertia of the column space is partitioned to identify polynomial components. when `catype = "SOCA"` or `catype = "SONSCA"`. When `catype = "DOCA"` or `catype = "DONSCA"`, the total inertia of both the row and column space is partitioned to give the polynomial components.

`Inner product` The inner product of the biplot coordinates for the two-dimensional plot.

`eccentricity` Value of ellipse eccentricity, the distance between its center and either of its two foci, It can be thought of as a measure of how much the conic section deviates from being circular.

`HL Axis 1` Value of ellipse semi-axis 1 for each row and column points.

`HL Axis 2` Value of ellipse semi-axis 2 for each row and column points.

Area	Ellipse area for each row and column points.
pvalcol	P-value for each row and column points.

Author(s)

Rosaria Lombardo and Eric J. Beh

References

Beh EJ and Lombardo R 2014 Correspondence Analysis: Theory, Practice and New Strategies. Wiley.
Lombardo R Beh EJ 2016 Variants of Simple Correspondence Analysis. The R Journal, 8 (2), 167–184.

Examples

```
data(asbestos)
resasbestos <- CAvariants(asbestos, catype = "DOCA")
print(resasbestos)
```

---

printwithaxes *Secondary printing function*

Description

The function is called from the main print function `print.CAvariants`. It adds the names to objects.

Usage

```
printwithaxes(x, thenames,digits=3)
```

Arguments

x	A matrix.
thenames	A character vector of the same length as x.
digits	The number of decimal places used for displaying the numerical summaries of the analysis. By default, digits = 3.

Author(s)

Rosaria Lombardo and Eric J. Beh

References

Beh EJ and Lombardo R 2014 Correspondence Analysis: Theory, Practice and New Strategies. Wiley.
Lombardo R Beh EJ 2016 Variants of Simple Correspondence Analysis. The R Journal, 8 (2), 167–184.

shopdataM *Two-way contingency table of Dutch shoplifting (1977-1978)*

Description

This two-way contingency table summarises, in part, the results of a survey of the Dutch Central Bureau of Statistics (Israels, 1987). The table considers a sample of 20819 men who were suspected of shoplifting in stores of the Netherlands between 1977 and 1978.

Usage

```
data(shopdataM)
```

Format

The format is:
row names [1:13] "clothing" "accessories" "tobacco" "stationary" ...
col names [1:9] "M12<" "M13" "M16" "M19" ...

References

Beh EJ and Lombardo R 2014 Correspondence Analysis: Theory, Practice and New Strategies. Wiley.
Israels A 1987 Eigenvalue Techniques for Qualitative Data. DSWO Press, Leiden.
Lombardo R Beh EJ 2016 Variants of Simple Correspondence Analysis. The R Journal, 8 (2), 167–184.

Examples

```
shopdataM <- structure(c(81, 66, 150, 667, 67, 24, 47, 430, 743, 132, 32,
197, 209, 138, 204, 340, 1409, 259, 272, 117, 637, 684, 408,
57, 547, 550, 304, 193, 229, 527, 258, 368, 98, 246, 116, 298,
61, 402, 454, 384, 149, 151, 84, 146, 141, 61, 40, 13, 71, 52,
138, 252, 942, 297, 313, 92, 251, 167, 193, 30, 16, 130, 111,
280, 624, 359, 109, 136, 36, 96, 67, 75, 11, 16, 31, 54, 200,
195, 178, 53, 121, 36, 48, 29, 50, 5, 6, 14, 41, 152, 88, 137,
68, 171, 37, 56, 27, 55, 17, 3, 11, 50, 211, 90, 45, 28, 145,
17, 41, 7, 29, 28, 8, 10, 28, 111, 34), .Dim =  c(13L,9L), .Dimnames = list(
c("clothing", "accessories", "tobacco", "stationary", "books",
"records", "household", "candy", "toys", "jewelry", "perfumes",
"hobby", "other"), c("M12<", "M13", "M16", "M19", "M25",
"M35", "M45", "M57", "M65+")))
dim(shopdataM)
```

`socabasic` *Singly, or one-way, ordered correspondence analysis: for an ordered column variable*

Description

This function is used by the main function `CAvariants` when the input parameter is `catype = "SOCA"`. It performs the hybrid decomposition of Pearson's ratios and computes the principal axes for the rows and polynomial axes for the columns. It also gives the coordinates, row and column weights, total inertia (based on Pearson's chi-squared statistic) and the rank of the matrix. It decomposes the inertia in terms of the column polynomial components.

Usage

```
socabasic(Xtable, mj)
```

Arguments

`Xtable`	The two-way contingency table.
`mj`	The set of ordered column scores. By default, `mj = c(1:ncol(Xtable))` (natural scores).

Author(s)

Rosaria Lombardo and Eric J. Beh

References

Beh EJ and Lombardo R 2014 Correspondence Analysis: Theory, Practice and New Strategies. Wiley.
Lombardo R Beh EJ 2016 Variants of Simple Correspondence Analysis. The R Journal, 8 (2), 167–184.
Lombardo R Beh EJ and Kroonenberg PM 2016 Modelling Trends in Ordered Correspondence Analysis Using Orthogonal Polynomials. Psychometrika, 81(2), 325–349.

Examples

```
data(asbestos)
mj <- c(1, 2, 3, 4, 5)
socabasic(asbestos, mj)
```

`sonscabasic` *Singly, or one-way, ordered non symmetrical correspondence analysis: for an ordered column predictor variable*

Description

This function is used by the main function `CAvariants` when the input parameter is `catype = "SONSCA"`. It performs the hybrid decomposition of the numerator of the Goodman-Kruskal tau index and implies an ordered (column) variable. It calculates the principal axes for the rows and polynomial axes for the columns, coordinates. It also calculates the row and column weights, inertia (based on the numerator of the tau index) and the rank of the matrix. It decomposes the inertia into column polynomial components.

Usage

```
sonscabasic(Xtable, mj)
```

Arguments

`Xtable`	The two-way contingency table.
`mj`	The set of ordered column scores. By default, `mj = c(1:ncol(Xtable))` (natural scores).

Author(s)

Rosaria Lombardo and Eric J. Beh

References

Beh EJ and Lombardo R 2014 Correspondence Analysis: Theory, Practice and New Strategies. Wiley.
Lombardo R Beh EJ 2016 Variants of Simple Correspondence Analysis. The R Journal, 8 (2), 167–184.
Lombardo R Beh EJ and Kroonenberg PM 2016 Modelling Trends in Ordered Correspondence Analysis Using Orthogonal Polynomials. Psychometrika, 81(2), 325-349.

Examples

```
data(asbestos)
mj<-c(1, 2, 3, 4, 5)
sonscabasic(asbestos, mj)
```

summary.CAvariants *Summary of numerical results from CA variants*

Description

This function prints a numerical summary of the results from any of the six variants of correspondence analysis. The input parameter is the name of the output of the main function CAvariants.

Usage

```
## S3 method for class 'CAvariants'
summary(object, printdims, digits, ...)
```

Arguments

object	The output of the main function CAvariants.
printdims	The number of dimensions that are used for summarising the numerical output of the analysis. By default, printdims = 2. the maximum number is equal to the rank of the table.
digits	The number of decimal places used for displaying the numerical summaries of the analysis. By default, digits = 3.
...	Further arguments passed to, or from, other functions.

Value

The value of output returned depends on the type of correspondence analysis that is performed.

Inertias
: The inertia values, their percentage contribution to the total inertia and the cumulative percent inertias for the row and column variables.

Generalised correlation matrix
: The matrix of generalised correlations when performing an ordered correspondence analysis, catype = "DOCA", catype = "DONSCA", catype = "SOCA" or catype = "SONSCA".

Row principal coordinates
: The row principal coordinates when catype = "CA", or catype = "NSCA".

Column principal coordinates
: The column principal coordinates when catype = "CA", or catype = "NSCA".

Row standard coordinates
: The row standard coordinates when catype = "CA", or catype = "NSCA".

Column standard coordinates
: The column standard coordinates when catype = "CA", or catype = "NSCA".

Row principal polynomial coordinates
: The row principal polynomial coordinates when catype = "DOCA", catype = "DONSCA", catype = "SOCA", or catype = "SONSCA".

Column principal polynomial coordinates
: The column principal coordinates when catype = "DOCA", or catype = "DONSCA".

`Row standard polynomial coordinates`
: The row standard polynomial coordinates when `catype` is `"DOCA"` or `"DONSCA"`.

`Column standard polynomial coordinates`
: The column standard polynomial coordinates when `catype = "DOCA"`, `catype = "DONSCA"`, `catype = "SOCA"`, or `catype = "SONSCA"`.

`Total inertia`
: The total inertia. For example, for non symmetrical correspondence analysis the numerator of the Goodman-Kruskal tau index, its C-statistic and p-value are returned.

`Polynomial components`
: The polynomial components of the total inertia and their p-values. The total inertia of the column space is partitioned to identify polynomial components. when `catype = "SOCA"` or `catype = "SONSCA"`. When `catype = "DOCA"` or `catype = "DONSCA"`, the total inertia of both the row and column space is partitioned to give the polynomial components.

`Inner product`
: The inner product of the biplot coordinates for the two-dimensional plot.

Author(s)

Rosaria Lombardo and Eric J. Beh

References

Beh EJ and Lombardo R 2014 Correspondence Analysis: Theory, Practice and New Strategies. Wiley.

Lombardo R Beh EJ 2016 Variants of Simple Correspondence Analysis. The R Journal, 8 (2), 167–184.

Lombardo R Beh EJ and Kroonenberg PM 2016 Modelling Trends in Ordered Correspondence Analysis Using Orthogonal Polynomials. Psychometrika, 81(2), 325–349.

Examples

```
asbestos <- matrix(c(310, 36, 0, 0, 212, 158, 9, 0, 21, 35, 17, 4, 25, 102,
49, 18, 7, 35, 51, 28), 4, 5, dimnames = list(c("none","grade1", "grade2", "grade3"),
c("0-9", "10-19", "20-29", "30-39", "40")))
risasbestos <- CAvariants(asbestos, catype = "DOCA", firstaxis = 1, lastaxis = 2)
summary(risasbestos)
```

---

trendplot *Trends of matrix rows and columns*

Description

This function portrays the row and column trends of the centred column profile matrix reconstructed by means of othogonal polynomials and/or principal axes.

Usage

```
trendplot(f, g, cex = 1, cex.lab = 0.8, main = " ", prop = 0.5,
posleg = "right", xlab = "First Axis",
ylab = "Second Axis")
```

Arguments

f	The row coordinates.
g	The column coordinates.
cex	The parameter for setting the size of character labels of points in graphical displays. By default, cex = 1.
cex.lab	The parameter for setting the size of character labels of axes in graphical displays. By default, cex.lab = 0.8.
main	The title of the graphical display.
prop	The scaling parameter for specifying the limits of the plotting area. By default, prop = 0.5.
posleg	The parameter for specifying the position of the legend in the graphical function trendplot. By default, pos = "right".
xlab	The parameter for setting the character label of the horizontal axis in graphical displays.
ylab	The parameter for setting the character label of the vertical axis in graphical displays.

Note

This function is called from the main plot function plot.CAvariants.

Author(s)

Rosaria Lombardo and Eric J. Beh

References

Beh EJ and Lombardo R 2014 Correspondence Analysis: Theory, Practice and New Strategies. John Wiley & Sons.
Lombardo R Beh EJ 2016 Variants of Simple Correspondence Analysis. The R Journal, 8 (2), 167–184.
Lombardo R Beh EJ and Kroonenberg PM 2016 Modelling Trends in Ordered Correspondence Analysis Using Orthogonal Polynomials. Psychometrika, 81(2), 325–349.

vcaellipse *Algebraic elliptical confidence regions for symmetrical variants of correspondence analysis*

Description

It produces elliptical confidence regions when symmetrical or ordered symmetrical correspondence analysis is performed. This function allows the analyst to superimpose confidence ellipses onto a graphical display when the input parameter `catype` of the main function `CAvariants` is set to `"CA"`, `"SOCA"` or `"DOCA"`. It is called internally from the main plot function `plot.CAvariants`. It uses the function `ellipse`.

Usage

```
vcaellipse(t.inertia, inertias, inertiapc, cord1, cord2, a, b, firstaxis=1,
lastaxis = 2, n, M = 2, Imass, Jmass)
```

Arguments

`t.inertia`	The total inertia of the two-way contingency table (Pearson's chi-squared or Goodman and Kruskal's index depends on the CA variant).
`inertias`	The explained inertia of each dimension.
`inertiapc`	The percentage of explained inertia for each dimension.
`cord1`	The row principal coordinates.
`cord2`	The column principal coordinates.
`a`	The row standard coordinates or, in case of the ordered variants of CA, the row standard polynomial coordinates.
`b`	The column standard coordinates or, in case of the ordered variants of CA, the column standard polynomial coordinates.
`firstaxis`	The horizontal polynomial, or principal, axis. By default, `firstaxis = 1`.
`lastaxis`	The vertical polynomial, or principal, axis. By default, `lastaxis = 2`.
`n`	The total number of observations.
`M`	The number of axes considered in determining the structure of the elliptical confidence regions.
`Imass`	The weight matrix of the row variable.
`Jmass`	The weight matrix of the column variable.

Details

The output values of this function.

Value

`eccentricity`	The eccentricity of the ellipses. This is the distance between the centre of the ellipse and its two foci, which can be thought of as a measure of how much the conic section deviates from being circular (when the region is perfectly circular, eccentricity is zero).
`HL Axis 1`	Value of the semi-major axis length for each row and column point.
`HL Axis 2`	Value of the semi-minor axis length for each row and column point.
`Area`	Area of the ellipse for each row and column point.
`pvalcol`	Approximate p-value for each of the row and column points.

Note

This function is called from the main plot function `plot.CAvariants` and is executed when in the main plot function the parameter `ell = TRUE`.

Author(s)

Rosaria Lombardo and Eric J Beh

References

Beh EJ 2010 Elliptical confidence regions for simple correspondence analysis. J. Stat. Plan. Inference 140, 2582–2588.
Beh EJ and Lombardo R 2014 Correspondence Analysis: Theory, Practice and New Strategies. Wiley.
Beh EJ Lombardo R 2015 Confidence regions and Approximate P-values for classical and non-symmetric correspondence analysis. Journal of Communications and Statistics, Theory and Methods. 44: 95–114.
Lombardo R Beh EJ 2016 Variants of Simple Correspondence Analysis. The R Journal, 8 (2), 167–184.

Index