Package 'CAvariants'

February 27, 2017

Type Package

Title Correspondence Analysis Variants

Version 3.4

Date 2017-01-20

Author Rosaria Lombardo and Eric J Beh

Maintainer Rosaria Lombardo <rosaria.lombardo@unina2.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

LazyData true

License GPL (> 2)

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

NeedsCompilation no

Repository CRAN

Date/Publication 2017-02-27 09:58:06

R topics documented:

`angle` *Angle*

Description

The **R** internal function `angle` is used to draw ellipses around points in graphical displays.

Details

This function is used in two other functions that are `caellipse` and `nscaellipse`.

References

Macdonald PDM 2002 Drawing an ellipse in Splus or R.
Available at www.math.mcmaster.ca/peter/s4m03/s4m03_0102/ellipse.html (accessed August 18, 2016).

`asbestos` *Selikoff's data, 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 of 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

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

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 the Pearson's ratio and compute principal axes, coordinates, weights of rows and columns, total inertia (equal to the 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. John Wiley & Sons. Lombardo R Beh EJ Variants of Simple Correspondence Analysis. The R Journal (accepted).

Examples

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

---

caellipse *Algebraic elliptical confidence regions*

Description

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

Usage

```
caellipse(Xtable, a1 = 1, a2 = 2, alpha = 0.05, cols = c(2, 4),
M = 2, cex = 0.8, cex.lab = 0.8,
mar = c(5, 4, 4, 2) + 0.1, prop = 0.8, Imass,
Jmass, a, b, g, fr, dmu, inertiapc,  plottype = "biplot",
biptype = "row", pos = 2, arrow = TRUE, length = 0, graphy = TRUE, ell = TRUE)
```

Arguments

Xtable	The two-way contingency table.
a1	The dimension reflected along the horizontal axis.
a2	The dimension reflected along the vertical axis.
alpha	The confidence level of the elliptical regions. By default, alpha = 0.05.
cols	The graphical parameter for setting the colours of the points in the graphical displays.
M	The number of axes used when drawing ellypse. By default, M = 2. Its maximum value is equal to the rank of the data matrix.
cex	The parameter for setting the size of character labels of points in graphical displays. By default, cex = 0.8.
cex.lab	The parameter for setting the size of character labels of axes in graphical displays. By default, cex.lab = 0.8.
mar	The parameter for setting the size of the plotting area.
prop	The scaling parameter for specifying the limits of the plotting area. By default, prop = 1.
Imass	The weight matrix of the row variable.
Jmass	The weight matrix of the column variable.
a	The row principal or polynomial axes.

b	The column principal or polynomial axes.
fr	The row coordinates.
g	The column coordinates.
dmu	The squared singular values or inertia of each axis.
inertiapc	The percentage of explained inertia.
pos	The parameter that specifies the position of point symbols in the graphical displays. By default, pos = 2.
plottype	The parameter for specifying the type of plot. By default, plottype = "biplot".
biptype	The parameter for specifying the type of biplot. By default, biptype = "row".
arrow	The parameter used for diplaying arrows in biplot in correspondence of points in principal coordinates. By default, arrow=TRUE.
length	The parameter used for setting the length of the arrow end in biplot. By default, length=0.
graphy	The flag parameter used for diplaying ellipses when using plot.CAvariants and not displaying when using print.CAvariants function. It is set as graphy = TRUE in plot.CAvariants and as graphy = FALSE in print.CAvariants.
ell	The flag parameter used for displaying ellipse. By default, ell = TRUE.

Details

The output values of this function.

Value

eccentricity	Value of ellipse eccentricity. This is the distance between the ellipse center and either of its two foci, which can be thought of as a measure of how much the conic section deviates from being circular (when it is equal to zero then the region becomes 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.

Note

This function is called from the main plot function plot.CAvariants and is executed when catype = "CA", catype = "SOCA" or catype = "DOCA", only if ell = TRUE.

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 Variants of Simple Correspondence Analysis. The R Journal (accepted).

`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", ellcomp = TRUE, Mell = 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. By default `firstaxis = 1`.
`lastaxis`	The vertical polynomial or principal axis. By default `lastaxis = 2`.
`catype`	The input parameter for specifying what variant of correspondence analysis is considered. By default, `catype = "CA"`. Other possible values are: `catype = "SOCA"`, `catype = "DOCA"`, `catype = "NSCA"`, `catype = "SONSCA"`, `catype = "DONSCA"`.
`ellcomp`	This input parameter ensures that the characteristics of the algebraic confidence ellipses are computed and stored. When `ellcomp = TRUE` (which is by default), the output includes the characteristics of the ellipses. The eccentricity of the confidence ellipses is summarised by the quantity `eccentricity`, this is the distance between the center and either of its two foci, which can be thought of as a measure of how much the conic section deviates from being circular (when it is equal to zero then the region becomes circular). The semi-major axis length of the ellipse for each row and column point is given by `HL Axis 1` while `HL Axis 2` gives the semi-minor axis length of the points along the second axis. The area of the ellipse for each row and column category is given by `Area` while the p-value of each category is defined by `P-value`.
`Mell`	The number of axes `Mell` considered in determining the structure of the elliptical confidence regions. By default, `Mell = min(nrow(Xtable), ncol(Xtable)) - 1`, i.e. the rank of the data matrix.
`alpha`	The confidence level of the elliptical regions. By default, `alpha = 0.05`.

Details

This function belongs to the object class called `cacorporateplus`

Value

Description of the output returned

`Xtable`	The two-way contingency table.
`rows`	The row number of the two-way contingency table.
`cols`	The column number of the two-way contingency table.
`r`	The rank of the two-way contingency table.
`rowlabels`	The label of the row variable.
`collabels`	The label of the column variable.
`Rprinccoord`	The row principal coordinates. When the input parameter `catype` is `"DOCA"`, `"SOCA"`, `"SONSCA"` or `"DONSCA"`, they are row principal polynomial coordinates.
`Cprinccoord`	The column principal coordinates. When the input parameter `catype` is `"DOCA"` or `"DONSCA"`, they are column principal polynomial coordinates.
`Rstdcoord`	The row standard coordinates. When the input parameter `catype` is `"DOCA"` or `"DONSCA"`, they are row standard polynomial coordinates.
`Cstdcoord`	The column standard coordinates. When the input parameter `catype` is `"DOCA"`, `"SOCA"`, `"SONSCA"` or `"DONSCA"`, they are column standard polynomial coordinates.
`tauden`	The tau denominator is given when the input parameter `catype` is `"NSCA"`, `"SONSCA"`, or `"DONSCA"`, otherwise it is a null value.
`tau`	The tau index is given when the input parameter `catype` is `"NSCA"`, `"SONSCA"`, or `"DONSCA"`, otherwise it is a null value.
`inertiasum`	The total inertia of the classical correspondence analysis when catype is `"CA"`, `"DOCA"` or `"SOCA"` (the Pearson's index), or the inertia of non symmetrical correspondence analysis when `catype` is `"NSCA"`, `"DONSCA"` or `"SONSCA"` (numerator of the Goodman-Kruskal tau index).
`inertias`	The associated inertia in absolute value and percentage, in the row space for each principal or polynomial axis.
`inertias2`	The associated 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.
`comps`	The polynomial components of inertia when the variables are ordered. The inertia of row and/or column space is partitioned in terms of polynomial components in ordered CA variants.
`catype`	The kind of correspondence analysis chosen.
`mj`	The ordered scores of a column variable. When `mj = NULL`, the natural ordered numbers are shown.

mi	The ordered scores of a row variable. When mi = NULL, the natural ordered numbers are shown.
pcc	The weighted centered column profile matrix.
Jmass	The weight matrix of the column variable.
Imass	The weight matrix of the row variable.
Trend	The inner product, Inner product, of the biplot coordinates (concerning the first two axes when firstaxis=1 and lastaxis=2)
Z	The generalized correlation matrix when catype is "SOCA", "DOCA" , "SONSCA", "DONSCA", but when catype is "CA", "NSCA", it gives again the inner product matrix of biplot coordinates.
ellcomp	The flag parameter, ellcomp, specifies that the characteristics of the confidence ellipses (eccentricity, semi-axis, area, p-values) are computed. By default, ellcomp = TRUE.
risell	When the input parameter, ellcomp, is set to ellcomp = TRUE, the output includes the characteristics risell of the confidence ellipses, the eccentricity of the confidence ellipses, risell$eccentricity, for each row and column point, the summary results, risell$row.summ and risell$col.summ, 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.
Mell	The number of axes Mell considered in determining the structure of the elliptical confidence regions. By default, Mell = min(nrow(Xtable), ncol(Xtable)) - 1, i.e. the rank of the data matrix.

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. This function belongs to the class cacorporateplus.

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 Variants of Simple Correspondence Analysis. The R Journal (accepted).

Examples

```
data(asbestos)
CAvariants(asbestos, catype = "CA", firstaxis = 1, lastaxis = 2)
CAvariants(asbestos, catype = "DOCA", firstaxis = 1,lastaxis = 2)
CAvariants(asbestos, catype = "DONSCA",firstaxis=1, lastaxis = 2, ellcomp = FALSE)
```

```
data(shopdataM)
CAvariants(shopdataM, catype = "NSCA", firstaxis = 1, lastaxis = 2)
CAvariants(shopdataM, catype = "SONSCA", firstaxis = 1, lastaxis = 2)
CAvariants(shopdataM, catype = "SOCA", firstaxis = 1, lastaxis = 2)
```

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

Description

This function allows the analyst to compute the contribution of the polynomial components to the inertia (chi-squared or tau). 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. John Wiley & Sons. Lombardo R Beh EJ Variants of Simple Correspondence Analysis. The R Journal (accepted).

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

Description

This function allows the analyst to compute the contribution of the polynomial components to the inertia (chi-squared or tau). The ordered variable should be the both row and column variables that are transformed by 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 component of inertia.

Note

This function belongs to the `R` object 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. John Wiley & Sons. Lombardo R Beh EJ Variants of Simple Correspondence Analysis. The R Journal (accepted).

`docabasic` *Doubly or two-way ordered correspondence analysis, it implies two ordered variables*

Description

This function is used in the main function `CAvariants` when the input parameter is `catype = "DOCA"`. It performs the bivariate moment decomposition of the Pearson's ratio, computes polynomial axes, coordinates, weights of rows and columns, total inertia (equal to the Pearson's index), the rank of the matrix. It allows one to decompose 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.
`mj`	The set of ordered column scores.

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. John Wiley & Sons. Lombardo R Beh EJ Variants of Simple Correspondence Analysis. The R Journal (accepted).

Examples

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

donscabasic *Doubly or two-way ordered non symmetrical correspondence analysis*

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 and implies two ordered variables. It computes polynomial axes, coordinates, weights of rows and columns, total inertia (equal to the numerator of the tau index) and the rank of the matrix. It allows the analyst to decompose 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.
mj	The set of ordered column scores.

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. John Wiley & Sons. Lombardo R Beh EJ Variants of Simple Correspondence Analysis. The R Journal (accepted).

Examples

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

ellipse *Ellipse*

Description

This function is called internally from the graphical functions `caellipse` and `nscaellipse`. It constructs the algebraic ellipses of confidence.

Arguments

`hlaxa`	The length of the horizontal ellipse semi-axis.
`hlaxb`	The length of the vertical ellipse semi-axis.
`xc`	The coordinate of the generic point on the horizontal axis.
`yc`	The coordinate of the generic point on the vertical axis.
`col`	The color of ellipses.

Note

This function is called from the secondary graphical function `caellipse` or `nscaellipse`, which is called from the main plot function `plot.CAvariants` and it can be executed for all variants of correspondence analysis.

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.
Macdonald PDM 2002 Drawing an ellipse in Splus or R.
Available at www.math.mcmaster.ca/peter/s4m03/s4m03_0102/ellipse.html (accessed August 18, 2016).

emerson.poly *Orthogonal polynomials*

Description

This function is called from the functions `docabasic, socabasic, sonscabasic` and `donscabasic`. It allows the analyst to compute the orthogonal polynomials of the ordered categorical variable. The number of the polynomials is equal to the variable category less one. The function computes the polynomial transformation of the ordered categorical variable.

Usage

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

Arguments

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

Value

Describe the value returned

B	the matrix of the orthogonal polynomials without the trivial polynomial.

Note

Note that the sum of the marginals of the ordered variables should be one.

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.
Emerson PL 1968 Numerical construction of orthogonal polynomials from a general recurrence formula. Biometrics, 24 (3), 695-701. Lombardo R Beh EJ Variants of Simple Correspondence Analysis. The R Journal (accepted).

Examples

```
emerson.poly(c(1,2,3,4,5), as.vector(c(.1,.2,.3,.2,.2)))
```

---

nscabasic *Non symmetrical two-way 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 rows and columns, total inertia (equal to the numerator of the tau index) and the rank of the matrix.

Usage

```
nscabasic(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. John Wiley & Sons. Lombardo R Beh EJ Variants of Simple Correspondence Analysis. The R Journal (accepted).

Examples

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

---

nscaellipse *Algebraic elliptical confidence regions in non symmetrical correspondence analysis*

Description

This function produces elliptical confidence regions when non symmetrical or ordered non symmetrical correspondence analysis is performed. It superimposes the confidence ellipses on to the graphical displays, when the input parameter catype of the main code CAvariants is equal to "NSCA", "SONSCA" or "DONSCA". It is called from the main plot function plot.CAvariants and uses the function ellipse.

Usage

```
nscaellipse(Xtable, a1 = 1, a2 = 2, alpha = 0.05, cols = c(2, 4),
M = 2, cex = 0.8, cex.lab = 0.8, mar = c(5, 4, 4, 2) + 0.1, prop = 0.8,  Imass, Jmass,
a, b, fr, g, dmu, tauden, inertiapc, plottype = "biplot",
biptype = "row", pos = 2, arrow = TRUE, length = 0, graphy = TRUE, ell = TRUE)
```

Arguments

Xtable	The two-way contingency table.
a1	The dimension reflected along the horizontal axis.
a2	The dimension reflected along the vertical axis.
alpha	The confidence level of the elliptical regions. By default, alpha = 0.05.
cols	The graphical parameter for setting the colours of the points in graphical displays.
M	The number of axes used when drawing ellypse. By default, M = 2. Its maximum value is equal to the rank of the data matrix.
cex	The parameter for setting the size of character labels of points in graphical displays. By default, cex = 0.8.
cex.lab	The parameter for setting the size of character labels of axes in graphical displays. By default, cex.lab = 0.8.
mar	The parameter for setting the size of the plotting area.
prop	The scaling parameter for specifying the limits of the plotting area. By default, prop = 1.
Imass	The weight matrix of the row variable.
Jmass	The weight matrix of the column variable.
a	The row principal or polynomial axes.
b	The column principal or polynomial axes.
fr	The row coordinates.
g	The column coordinates.
dmu	The squared singular values or inertia of each axis.
tauden	The denominator of tau index.
inertiapc	The percentage of explained inertia.
plottype	The kind of graphical display. By default, plottype = "biplot".
biptype	The parameter for specifying the type of biplot. By default, biptype = "row".
pos	The parameter that specifies the position of the point symbols in the graphical displays. By default, pos = 2.
arrow	The parameter for diplaying arrows in biplot in correspondence of points in principal coordinates. By default, arrow = TRUE.
length	The parameter for setting the length of the arrow end in biplot. By default, length = 0.
graphy	The parameter for diplaying ellipses in plots or biplots. By default, graphy = TRUE.
ell	The flag parameter used for displaying ellipse. By default, ell = TRUE.

Details

The output values of this function

Value

eccentricity	Value of ellipse eccentricity. This is the distance between the ellipse center and either of its two foci, which can be thought of as a measure of how much the conic section deviates from being circular (when it is equal to zero then the region becomes 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.

Note

This function is called from the main plot function `plot.CAvariants` and is executed when `catype = "NSCA"`, `catype = "SONSCA"` or `catype = "DONSCA"`, only if `ell = TRUE`.

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 Variants of Simple Correspondence Analysis. The R Journal (accepted).

---

plot.CAvariants *Main plot function*

Description

This function allows the analyst to produce the suitable graphical displays with respect to the six variants of correspondence analysis. In particular when `plottype = "classic"`, it produces classical graphical displays for `catype = "CA"` and `catype = "NSCA"`, where the row and column variables are graphically depicted in principal coordinates.
When we set `plottype = "biplot"`, it produces biplot graphical displays, or polynomial biplots in case of ordered analysis. Notice that for ordered analysis only polynomial biplots are suitable. In particular, for the singly ordered variants only row isometric polynomial biplots make sense, as we assume that the ordered variable is the column variable (the column coordinates are standard polynomial coordinates and the row coordinates are principal polynomial coordinates). When the input parameter `catype` is equal to `catype = "SOCA"`, `catype = "DOCA"`, `catype = "SONSCA"` or `catype = "DONSCA"`, then the input parameter `plottype` should be equal to `plottype = "biplot"`, if `biptype = "row"` , it will give back a row isometric polynomial biplot.

Usage

```
## S3 method for class 'CAvariants'
plot(x, firstaxis = 1, lastaxis = 2, cex = 0.8, cex.lab = 0.8,
prop = 1,  plottype = "biplot", biptype = "row",
scaleplot = 1, posleg = "topleft", pos = 2, ell = FALSE, Mell = x$Mell, alpha = 0.05,...)
```

Arguments

x	The name of the output object, for example say res, used with the main function CAvariants.
firstaxis	The horizontal polynomial or principal axis, firstaxis. By default, firstaxis = 1.
lastaxis	The vertical polynomial or principal axis, lastaxis. By default, lastaxis = 2.
cex	The size of characters, cex, displayed on the correspondence plot or biplot. By default, cex = 0.8.
cex.lab	The parameter cex.lab that specifies the size of character labels of axes in graphical displays. By default, cex.lab = 0.8.
prop	The scaling parameter for specifying the limits of the plotting area. By default, prop = 1.
plottype	The type of graphical display required (either a classical correspondence plot or a biplot). The user can look at a classical correspondence plot by defining the input parameter plottype = "classic". When plottype = "biplot", it produces biplot graphical displays, or polynomial biplots in case of an ordered analysis. Note that for ordered analysis only polynomial biplots are suitable. In particular for the singly ordered variants, only row isometric polynomial biplots make sense, as we assume that the ordered variable is the column variable (the column coordinates are standard polynomial coordinates and the row coordinates are principal polynomial coordinates). By default, plottype = "biplot".
biptype	For a biplot, one may specify that it be a row-isometric biplot (biptype = "row") or a column-isometric biplot (biptype = "column"). This feature is available for the nominal symmetrical and the non symmetrical correspondence analyses. By default, a row-isometric biplot, biptype = "row", is produced.
scaleplot	The parameter for scaling the biplot coordinates, scaleplot, originally proposed in Section 2.3.1 of Gower et al. (2011) and described on page 135 of Beh and Lombardo (2014). By default, scaleplot = 1.
posleg	The parameter posleg for specifying the position of the legend when portraying trends of ordered categories in ordered variants of correspondence analysis. By default, posleg = "topleft".
pos	The parameter for specifying the position of point symbols in the graphical displays. By default, pos = 2.
ell	The logical parameter, ell which specifies whether algebraic confidence ellipses are to be included in the plot or not. Setting the input parameter to ell = TRUE will allow the user to assess the statistical significance of each category to the association between the variables. The ellipses will be included when the plot is constructed using principal coordinates (being either row and

column principal coordinates or row and column principal polynomial coordinates). By default, this input parameter is set to `ell = FALSE`. See also the input parameter `ellcomp` of the function `CAvariants` for a description of the numeric characteristics of the confidence ellipses (eccentricity, area, etc.), as well as the input parameter `ellprint` of the function `print.CAvariants` for getting a print of these parameters.

`Mell` The number of axes `Mell` considered when portraying the elliptical confidence regions.
By default, it is equal to `Mell = min(nrow(Xtable), ncol(Xtable)) - 1`, i.e. the rank of the data matrix. This parameter is identical to the input parameter `Mell` of the function `CAvariants`.

`alpha` The confidence level of the elliptical regions. By default, `alpha = 0.05`.

`...` Further arguments passed to or from other methods.

Details

It produces classical and biplot graphical displays. Further when `catype` is equal to `"DOCA", "SOCA", "DONSCA"` or `"SONSCA"`, the trend of row and column variables after the reconstruction of column profiles by polynomials is portrayed.
For classical biplot displays, it superimposes on it algebraic ellipses of confidence. It uses the secondary plot functions `caellipse` or `nscaellipse`, depending on the input parameter `catype`.

Note

For classical graphical displays, both sets of coordinates are defined using principal coordinates (see Greenacre's terminology). In biplot graphical displays, one set of coordinates is standard and the other is principal. When the analysis is ordered, it makes sense only biplot. One set of coordinates consists of standard polynomial coordinates and the other one is of 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. John Wiley & Sons. Gower J, Lubbe S, and le Roux, N 2011 Understanding Biplots. John Wiley & Sons. Lombardo R Beh EJ Variants of Simple Correspondence Analysis. The R Journal (accepted).

Examples

```
data(asbestos)
risasbestos<-CAvariants(asbestos, catype = "DOCA", firstaxis=1, lastaxis=2)
plot(risasbestos, plotype = "biplot", biptype = "row")
plot(risasbestos, plotype = "biplot", biptype = "row", ell = TRUE)
```

plotone *Secondary plot function*

Description

This function produces a classical plot when the input parameter of the main code `CAvariants` has been set as `plottype = "classic"`, or it gives a biplot when `plottype = "biplot"`. This function allows the analyst to plot row and column points in graphical displays, using symbols and labels for each point. It is called internally from the main graphical function `plot.CAvariants`.

Usage

```
plotone(a1, a2, plottype, things, nthings, nvars, Thingcoord, Varcoord,
inertiapc, thinggroup, thinggrlab, vargroup, vargrlab, thinglabels, varlabels,
picsize, cex = 0.8, cex.lab = 0.8, type , catype , pos = 2)
```

Arguments

`a1`	The dimension reflected along the horizontal (polynomial or principal) axis in the graphical display.
`a2`	The dimension reflected along the vertical (polynomial or principal) axis in the graphical display.
`plottype`	The paramater for specifying what kind of graphical display is required. By default, `plottype = "classic"`, otherwise `plottype = "biplot"`.
`things`	The input parameter for specifying what variable is in principal coordinates. Its value is `"row"` or `"column"` in biplots `"rows&columns"` in classical plots.
`nthings`	The number of row categories.
`nvars`	The number of column categories.
`Thingcoord`	The coordinates of the row variable.
`Varcoord`	The coordinates of the column variable.
`inertiapc`	The percentage of explained inertia for each axis.
`thinggroup`	It is a list of two objects, the first is the vector of the ordered natural scores of the row variable, the second object is a vector of one whose length is equal to the row number.
`thinggrlab`	The set of plotting symbols and colours of the row variable.
`vargroup`	It is a list of two objects, the first is the vector of the ordered natural scores of the column variable, the second object is a vector of one whose length is equal to the column number.
`vargrlab`	The set of plotting symbols and colours of the column variable.
`thinglabels`	The labels of the row variable.
`varlabels`	The labels of the column variable.
`picsize`	The graphical parameter for specifying the size of the plotting area.

cex	The parameter for setting the size of character labels of points in graphical displays. By default, cex = 0.8.
cex.lab	The parameter that specifies the size of character labels of axes in graphical displays. By default, cex.lab = 0.8.
type	The graphical parameter that specifies the type of line to be drawn.
catype	The input parameter specifying what variant of correspondence analysis is to be performed.
pos	The parameter for specifying the position of point symbols in the graphical displays. By default, pos = 2.

Note

In classical graphical displays, both sets of coordinates are principal coordinates (see Greenacre's terminology). In biplot graphical displays, one set of coordinates is standard and the other is principal. When the analysis is ordered, one set of coordinates consists of standard polynomial coordinates and the other one is of principal polynomial coordinates. This function is internally called by the main graphical 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 Variants of Simple Correspondence Analysis. The R Journal (accepted).

---

print.CAvariants *Main printing function*

Description

This function prints the results of any of the specified 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'
print(x, printdims = 2, ellprint = TRUE, Mell = min(nrow(x$Xtable), ncol(x$Xtable)) - 1,
 alpha = 0.05, digits = 3,...)
```

Arguments

x	The name of the output object, for example say res, used with the main function CAvariants.
printdims	The number of dimensions, printdims, that are used to generate the correspondence plot, or biplot, and for summarising the numerical output of the analysis. By default, printdims = 2.
ellprint	The flag parameter, ellprint, allows that the characteristics of the confidence ellipses (eccentricity, semi-axis, area, p-values) are displayed. By default, ellprint = TRUE.
Mell	The number of axes Mell used for the construction of the confidence ellipses. By default, it is equal to its maximum value, Mell = min(nrow(x$Xtable), ncol(x$Xtable)) - 1, i.e. the rank of the data matrix. This input parameter is identical to the parameter Mell of both the functions CAvariants and plot.CAvariants.
alpha	The level of significance used for the construction of the elliptical regions. By default, alpha = 0.05.
digits	The minimum number of decimal places, digits, used for displaying the numerical summaries of the analysis. By default, digits = 3.
...	Further arguments passed to or from other methods.

Details

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

Value

The value of output returned depends on the kind of correspondence analysis performed

Xtable	The two-way contingency table.
Row weights: Imass	The row weight matrix. These weights depend on the type of analysis performed.
Column weights: Jmass	The column weight matrix. These weights are equal to the data column margins for all types of analysis performed.
Total inertia	The total inertia of the analysis performed. For example, when considering the variants of non symmetrical correspondence analysis, the numerator of the Goodman-Kruskal tau index, the associated C-statistic and its p-value is produced.
Inertias	The inertia values, their percentage contribution to the total inertia and the cumulative percent inertias of the row and column space. When performing an ordered correspondence analysis, this output summary describes both the row and column spaces for each principal or polynomial axis. When catype is CA or NSCA, the associated inertia values in the row and column spaces are identical.

`Generalized correlation matrix`
: The generalized correlation matrix when performing an ordered correspondence analysis,
`catype` should be `"DOCA"`, `"DONSCA"`, `"SOCA"` or `"SONSCA"`.

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

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

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

`Column standard coordinates`
: The column standard coordinates when `catype` is `"CA"` or `"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, i.e. standard polynomial coordinates for the row categories, when performing a doubly ordered correspondence analysis.

`Column standard polynomial coordinates`
: The column standard polynomial coordinates, i.e. standard polynomial coordinates for the column categories, when performing an ordered correspondence analysis.

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

`Column distances from the origin of the plot`
: The 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` is `"SOCA"` or `"SONSCA"`. When `catype`
is `"DOCA"` or `"DONSCA"`, the total inertia of both the row and column space is partitioned to identify of polynomial components.

`Inner product`
: The inner product of the biplot coordinates (concerning the first two axes when `firstaxis = 1` and `lastaxis = 2`).

`ellprint`
: When the input flag parameter is `ellprint = TRUE`, then the print includes the eccentricity of the confidence ellipses, the semi-major axis length of the ellipse for each row and column point, `HL Axis 1`, the semi-minor axis length for the ellipse for each row and column point, `HL Axis 2`, the area of the ellipse for each row and column point, `Area` and the p-value for each row and column point, `P-value`, see also the parameter `ellcomp` of the function `CAvariants` for a detailed description of these parameters.

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.

Examples

```
data(asbestos)
risasbestos<-CAvariants(asbestos, catype = "DOCA", firstaxis = 1, lastaxis = 2)
print(risasbestos)
```

---

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`	An R object.
`thenames`	A character vector of up to the same length as `x`.
`digits`	The minimum number of significant digits to be printed in values. By default, `digits = 3`.

Note

It is called from `print.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.

shopdataM *Two-way contingency table*

Description

This two-way contingency table summarises, in part, the results of a survey of the Dutch Central Bureau of Statistics (Isra\"els, 1987). The table considers a sample of 20819 men who were suspected of shoplifting in Dutch stores 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. John Wiley & Sons.
Isra\"els A 1987 Eigenvalue Techniques for Qualitative Data. DSWO Press, Leiden.

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, it implies an ordered (column) variable*

Description

This function is used in the main function `CAvariants` when the input parameter is `catype = "SOCA"`. It performs the hybrid decomposition of the Pearson's ratio and computes principal axes for rows and polynomial axes for columns, coordinates, weights of rows and columns, total inertia (equal to the Pearson's index) and the rank of the matrix. It allows the analyst to decompose 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.

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. John Wiley & Sons. Lombardo R Beh EJ Variants of Simple Correspondence Analysis. The R Journal (accepted).

Examples

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

sonscabasic *Singly or one-way ordered non symmetrical correspondence analysis*

Description

This function is used in 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 compute the principal axes for rows and polynomial axes for columns, coordinates, weights of rows and columns, inertia (equal to the numerator of the tau index), the rank of the matrix. It allows the analyst to decompose the inertia in column polynomial components.

Usage

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

Arguments

Xtable	The two-way contingency table.
mj	The set of ordered column scores.

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. John Wiley & Sons. Lombardo R Beh EJ Variants of Simple Correspondence Analysis. The R Journal (accepted).

Examples

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

summary.CAvariants *Summary of results of CA variants analysis*

Description

This function prints summary results of any of the specified 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 dimension number in printing. By default, printdims = 3.
digits	The minimum number of significant digits to be printed in values. By default, digits = 3.
...	Further arguments passed to or from other methods.

Value

The value of output returned depends on the kind of correspondence analysis performed.

Inertias	The inertia value and its percentage, in the row and column space for each principal or polynomial axis when catype is "DOCA", "DONSCA", "SOCA" or "SONSCA". When catype is "CA" or "NSCA", the associated inertia values in the row and column spaces are the same.
Generalized correlation matrix	The generalized correlation matrix when catype is "DOCA", "DONSCA", "SOCA" or "SONSCA".
Row principal coordinates	The row principal coordinates when catype is "CA" or "NSCA".
Column principal coordinates	The column principal coordinates when catype is "CA" or "NSCA".
Row standard coordinates	The row standard coordinates when catype is "CA" or "NSCA".
Column standard coordinates	The column standard coordinates when catype is "CA" or "NSCA".
Row principal polynomial coordinates	The row principal polynomial coordinates when catype is "DOCA", "DONSCA", "SOCA" or "SONSCA".
Column principal polynomial coordinates	The column principal coordinates when catype is "DOCA","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` is `"DOCA"`, `"DONSCA"`, `"SOCA"`, `"SONSCA"`.

`Total inertia`
: The total inertia. For example in case of non symmetrical correspondence analysis, it gives the numerator of the Goodman-Kruskal tau index, the associated C-statistic and its p-value.

`Polynomial components`
: The polynomial components of inertia and their p-values. The inertia of the column space is partitioned in terms of polynomial components when `catype` is `"SOCA"` or `"SONSCA"`. The inertia of the row and column space is partitioned in terms of polynomial components when `catype` is `"DOCA"` or `"DONSCA"`.

`Inner product`
: The inner product of coordinates.

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 Variants of Simple Correspondence Analysis. The R Journal (accepted).

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 = "topleft", xlab = "First Principal Axis",
ylab = "Second Principal 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 = "topleft".
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 Variants of Simple Correspondence Analysis. The R Journal (accepted).

Index