Package 'asymmetry'

February 16, 2017

Type Package

Title Visualizing Asymmetric Data

Version 1.2

Date 2017-02-07

Author Berrie Zielman

Maintainer Berrie Zielman <berrie.zielman@gmail.com>

Description Models and methods for the visualization for asymmetric data. A matrix is asymmetric if the number of rows equals the number of columns, and these rows and columns refer to the same set of objects. An example is a student migration table, where the rows correspond to the countries of origin of the students and the columns to the destination countries. The slide-vector model, a heat map, and the decomposition of asymmetry are provided for the analysis of these tables.

License GPL (>= 2)

Imports rJava, gplots, stats, methods, smacof

Suggests knitr, rmarkdown, RColorBrewer

VignetteBuilder knitr

NeedsCompilation yes

SystemRequirements Java JDK 1.2 or higher (for JRI/REngine JDK 1.4 or higher), GNU make

Repository CRAN

Date/Publication 2017-02-16 22:38:23

R topics documented:

`asymmetry-package` *Visualizing Asymmetric Data*

Description

Models and methods for the visualization for asymmetric data. A matrix is asymmetric if the number of rows equals the number of columns, and these rows and columns refer to the same set of objects. An example is a student migration table, where the rows correspond to the countries of origin of the students and the columns to the destination countries. The slide-vector model, a heat map, and the decomposition of asymmetry are provided for the analysis of these tables.

Author(s)

Berrie Zielman

Maintainer: Berrie Zielman <berrie.zielman@gmail.com>

References

Zielman, B., and Heiser, W. J. (1993), The analysis of asymmetry by a slide-vector, Psychometrika, 58, 101-114.

`asymscal` *asymscal*

Description

This function fits the asymscal model proposed by Young (1975). Asymmetry in the data arises by weighting the dimensions of a multidimensional scaling configuration. When a subject compares object *i* to *j* he or she may use different weights when comparing object *j* to *i*.

Usage

```
asymscal(data, ndim = 2, itmax = 10000, eps = 1e-10)
```

Arguments

data	The data
ndim	The number of dimensions
itmax	The maximum number of iterations
eps	Convergence criterion for Stress

Details

This function exploits a connection between the INDSCAL model and the asymscal model. This method inherits the methods for plotting an printing from the smacofIndDiff in the smacof package. Basically, the asymscal takes two steps. First, this function sets up the appropriate dissimilarity and missing data structure for a three-way multidimensional scaling model, then a call to the method smacofIndDiff in the imported package smacof is made. After correcting for the normalization applied to the data by smacofIndDiff, the results can be displayed and plotted by the methods in the package smacof.

Value

delta	Observed dissimilarities
obsdiss	List of observed dissimilarities, normalized
confdiss	List of configuration dissimilarities
conf	List of matrices of final configurations
gspace	Joint configurations aka group stimulus space
cweights	Configuration weights
stress	Stress-1 value
resmat	Matrix with squared residuals
rss	Residual sum-of-squares
spp	Stress per point
ndim	Number of dimensions
model	Type of smacof model
niter	Number of iterations
nobj	Number of objects

Warning

Although my experience with fitting this model in the majorization framework is very limited, it appears to converge slowly. I recommend keeping the convergence criteria in the example

Note

The original algorithm for fitting the asymscal model fits squared distances. This function fits distances and not squared distances. The configuration matrix is normalized to one.

Author(s)

Berrie Zielman

References

Young, F. W. (1975). An asymmetric Euclidean model for multi-process asymmetric data. Paper presented at the U.S.-Japan Seminar on Multidimensional scaling, San Diego, U.S.A.

Examples

```
## Not run:
data("asymscalexample")
t<-asymscal(data,ndim=2,itmax=10000,eps=1e-10)
t$cweights
round(t$cweights,3)
plot(t, plot.type = "confplot")
plot(t, plot.type = "bubbleplot")
plot(t, plot.type = "stressplot")

## End(Not run)
```

---

asymscalexample *Asymscal Example Data*

Description

This is an artificial dataset. The data are distances from a two-dimensional model, and because of this construction the asymscal model fit this data exactly. In addition, two rows of this matrix have weights different from (1,1). The fifth subject has weights (1.35,.25), and the 15th subject has weights (1.65,.425).

Usage

```
data("asymscalexample")
```

Format

A matrix with 15 rows and 15 columns.

Examples

```
data(asymscalexample)
## maybe str(asymscalexample) ; plot(asymscalexample) ...
```

`decomposition` *Decompose an asymmetric matrix*

Description

The decomposition of an asymmetric matrix into a symmetric matrix and a skew-symmetric matrix is an elementary result from mathematics that is the cornerstone of this package. The decomposition into a skew-symmetric and a symmetric component is written as: $Q = S + A$, where Q is an asymmetric matrix, S is a symmetric matrix, and A is a skew-symmetric matrix. This decomposition provides a justification for separate analyses of S and A. This decompositon is a usefull tool for data analysis, and a second application is to the study of an asymmetric matrix of residuals, obtaind after fitting a model.

Usage

```
decomposition(x)
```

Arguments

`x`	An asymmetric matrix

Value

`S`	The symmetric part of the matrix
`A`	The skew-symmetric part of the matrix
`linear`	The linear part of the skew-symmetric matrix

Author(s)

Berrie Zielman

Examples

```
data("Englishtowns")
Q <- decomposition(Englishtowns)
# the skew-symmetric part
Q$A
```

Englishtowns *Distance matrix of eight english towns*

Description

A data matrix with 8 rows and 8 columns. The data are a distance matrix between eight English towns. The data are not symmetric because a linear skew-symmetric matrix is added to the true distances

Usage

```
data("Englishtowns")
```

References

Constantine, A.G. & Gower, J.C. (1978). Graphical Representation of Asymmetric Matrices. Appl. Statist, 27, 297-304.

Examples

```
data(Englishtowns)
```

hmap *Heat map for skew-symmetric data*

Description

This heatmap displays the values of a skew-symmetric matrix by colors. The option `dominance` orders the rows and columns of the matrix in such a way that the values in the uppertriangle are positive and the values in the lower triangle are negative. The order is calculated from the row-sums of the signs obtained from the skew-symmetric matrix.

Usage

```
hmap(x, dominance = FALSE, ...)
```

Arguments

`x`	A skew-symmetric matrix, an asymmetric matrix, or an object of class `decomposition` If an asymmetric matrix is given, the skew-symmetrix part is computed.
`dominance`	If true the signs of the skew-symmetric matrix are shown in the heatmap, if set to false the values in this matrix are shown.
`...`	Further plot arguments: see `heatmap.2` for detailed information.

Author(s)

Berrie Zielman

Examples

```
data(studentmigration)
hmap(studentmigration,dominance=TRUE, col=c("red","white","blue"))
```

`plot.slidevector` *Configuration Plot for the Slide-vector Model*

Description

Method for a two-dimensional plot of the model. Available rownames are plotted as labels above the points. The slide-vector is shown as an arrow.

Usage

```
## S3 method for class 'slidevector'
plot(x, plot.dim = c(1,2), yplus = 0, xlab, ylab, ...)
```

Arguments

`x`	Object of class `slidevector`
`plot.dim`	A vector with dimensions to be plotted
`yplus`	parameter to adjust the vertical position of the label
`xlab`	Label of x-axis.
`ylab`	Label of y-axis.
`...`	Further plot arguments: see `plot` for detailed information.

Examples

```
## 2D plot for the slide-vector model
dis <- matrix(c(1,2,3,4,5,6,2,8,9,3), nrow = 5, ncol = 2)
a <- rbind(dis,dis+1.5)
test <- as.matrix(dist(a))[1:5,6:10]
v <- slidevector(test, dim = 2,itmax = 250, eps = .001)
plot(v)
```

print.slidevector *Print Method for the Slide-vector Model*

Description

Print method for objects of class slidevector.

Usage

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

Arguments

x	Object of class slidevector
...	Further arguments

slidevector *The slide-vector model*

Description

The slide vector model is a multidimensional scaling model for asymmetric data. A distance model is fitted to the symmetric part of the data whereas the asymmetric part of the data is represented by projections of the coordinates onto the slide-vector. The slide-vector points in the direction of large asymmetries in the data. The interpretation of asymmetry in this model is aided by the use of projections of points onto the slide-vector. The distance from i to j is larger if the point i has a higher projection onto the slide-vector than the distance from j to i. If the line connecting two points is perpendicular to the slide-vector the difference between the two projections is zero. In this case the distance between the two points is symmetric. The algorithm for fitting this model is derived from the majorization approach to multidimensional scaling.

Usage

```
slidevector(data, dim = 2, verbose = FALSE, itmax = 125, eps = 1e-12)
```

Arguments

data	An asymmetric matrix
dim	The number of dimensions for this model
verbose	Print the history of iterations
itmax	The maximum number of iterations
eps	The convergence criterion for the algorithm

Details

The slide-vector model is a special case of the unfolding model. Therefore, the algorithm for fitting this model is a constrained unfolding model. The coordinates of the objects are calculated by minimizing a least squares loss function. This loss function is called stress in the multidimensional scaling literature. The stress is minimized by a version of the SMACOF algorithm. The main output are the configuration of points and the slide-vector.

Value

`ndim`	the number of dimensions
`stress`	the raw stress for this model
`confi`	returns the configuration matrix of this multidimensional scaling model
`niter`	the number of iterations for the algorithm to converge
`nobs`	the number of observations
`resid`	A matrix with raw residuals
`slvec`	Coordinates of the slide-vector

Author(s)

Berrie Zielman

References

Zielman, B., and Heiser, W. J. (1993), The analysis of asymmetry by a slide-vector, Psychometrika, 58, 101-114.

See Also

`plot.slidevector`

Examples

```
## asymmetric distances between English towns
data(Englishtowns)
v <- slidevector(Englishtowns,dim = 2,itmax = 250, eps = .001)
plot(v)
```

`studentmigration` *Student Migration in the Erasmus Program*

Description

The table lists the home and destination country of 268.142 students in the academic year 2012-2013. The 33 rows of this table refer to the home country whereas the 33 columns refer to the destination countries. The table gives the number of inbound and outbound students between every pair of countries. The entries in the table are read as follows: 32 students from Bulgaria studied in The Netherlands, 18 students from the Netherlands studied in Bulgaria. Macedonia (MK) was excluded from the published table because only one student from Macedonia studie abroad and this country did not receive any students.

Usage

```
data(studentmigration)
```

Format

A matrix of 33 rows by 33 columns

Details

The Erasmus program is a student exchange program from the European Union. Three million students had taken part since the start of the program in 1987. To join the program a student has study at least three months or do an internship of at least two months in another country. The 2-letter codes shown below are supplied by the ISO (International Organization for Standardization). Country codes are given here: Countrycodes

Note

Macedonia has been removed from this table because only one student from this country participated in the program.

Source

http://ec.europa.eu/education/library/statistics/ay-12-13/annex-2_en.pdf

Examples

```
data(studentmigration)
hmap(studentmigration)
```

summary.decomposition *Summary method of the decomposition*

Description

Prints a decomposition of the sum of squares of an asymmetric matrix. The first column gives the sum of squares, and the second column gives the percentages of the two components. This decomposition can be applied to data, but also to a matrix of residuals obtained from a fitted model.

Usage

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

Arguments

object	An object of class decomposition
...	further parameters

Author(s)

Berrie Zielman

Examples

```
data(Englishtowns)
q <- decomposition(Englishtowns)
summary(q)
```

summary.slidevector *Summary method for the slide-vector model*

Description

This function extracts the configuration matrix and the slide-vector from the object.

Usage

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

Arguments

object	An object of class slidevector
...	Ignored

Examples

```
data(Englishtowns)
v <- slidevector(Englishtowns,dim = 2,itmax = 250, eps = .001)
summary(v)
```

Index