Package ‘RSDA’

November 10, 2023

Type Package

Title R to Symbolic Data Analysis

Version 3.2.1

Date 2023-11-09

Description Symbolic Data Analysis (SDA) was proposed by professor Edwin Diday in 1987, the main purpose of SDA is to substitute the set of rows (cases) in the data table for a concept (second order statistical unit). This package implements, to the symbolic case, certain techniques of automatic classification, as well as some linear models.

License GPL (>= 2)

Encoding UTF-8

Depends R (>= 3.4)

URL https://oldemarrodriguez.com/

Suggests testthat (>= 2.1.0), knitr, rmarkdown

RoxygenNote 7.2.3

Imports vctrs (>= 0.2.4), dplyr (>= 0.8.5), forcats, scales, stringr, rlang (>= 0.4.5), purrr, magrittr, tidyselect, tibble (>= 3.0.0), stats, RJSONIO, XML, ggplot2, ggpolypath, reshape, glmnet, FactoMineR, princurve, nloptr, sqldf, randomcoloR, kknn, e1071, gbm, randomForest, rpart, neuralnet, umap, xtable, plotly, ggrepel

VignetteBuilder knitr

LazyData true

NeedsCompilation no

Author Oldemar Rodriguez [aut, cre],
Jose Emmanuel Chacon [cph],
Carlos Aguero [cph],
Jorge Arce [cph]

Maintainer Oldemar Rodriguez <oldemar.rodriguez@ucr.ac.cr>

Repository CRAN

Date/Publication 2023-11-10 14:53:24 UTC

R topics documented:

`abalone` *SODAS XML data file.*

Description

Example of SODAS XML data file converted in a CSV file in RSDA format.

Usage

```
data(abalone)
```

Format

An object of class `symbolic_tbl` (inherits from `tbl_df`, `tbl`, `data.frame`) with 24 rows and 7 columns.

Source

http://www.info.fundp.ac.be/asso/sodaslink.htm

References

Bock H-H. and Diday E. (eds.) (2000).Analysis of Symbolic Data. Exploratory methods for extracting statistical information fromcomplex data. Springer, Germany.

Examples

```
data(abalone)
res <- sym.pca(abalone, 'centers')
plot(res, choix = "ind")
plot(res, choix = "var")
```

`as.data.frame.symbolic_histogram`
a data.frame

Description

a data.frame

Usage

```
## S3 method for class 'symbolic_histogram'
as.data.frame(x, ...)
```

Arguments

`x`	.....
`...`	...

as.data.frame.symbolic_interval

convertir a data.frame

Description

convertir a data.frame

Usage

```
## S3 method for class 'symbolic_interval'
as.data.frame(x, ...)
```

Arguments

x	a symbolic interval vector
...	further arguments passed to or from other methods.

as.data.frame.symbolic_modal

Extract values

Description

Extract values

Usage

```
## S3 method for class 'symbolic_modal'
as.data.frame(x, ...)
```

Arguments

x	An object to be converted
...	Further arguments to be passed from or to other methods.

as.data.frame.symbolic_set *convertir a data.frame*

Description

convertir a data.frame

Usage

```
## S3 method for class 'symbolic_set'
as.data.frame(x, ...)
```

Arguments

x	a symbolic interval vector
...	further arguments passed to or from other methods.

calc.burt.sym *Burt Matrix*

Description

Burt Matrix

Usage

```
calc.burt.sym(sym.data, pos.var)
```

Arguments

sym.data	ddd
pos.var	ddd

`calculate.quantils.RSDA` *quantiles.RSDA*

Description

quantiles.RSDA

Usage

```
calculate.quantils.RSDA(histogram.RSDA, num.quantils)
```

Arguments

`histogram.RSDA`	A histogram
`num.quantils`	Number of quantiles

Value

Quantiles of a Histogram

`Cardiological` *Cardiological data example*

Description

Cardiological interval data example.

Usage

```
data(Cardiological)
```

Format

An object of class `symbolic_tbl` (inherits from `tbl_df`, `tbl`, `data.frame`) with 11 rows and 3 columns.

References

Billard L. and Diday E. (2006).Symbolic data analysis: Conceptual statistics and data mining. Wiley, Chichester.

Examples

```
data(Cardiological)
res.cm <- sym.lm(formula = Pulse~Syst+Diast, sym.data = Cardiological, method = 'cm')
pred.cm <- sym.predict(res.cm, Cardiological)
RMSE.L(Cardiological$Pulse, pred.cm$Fitted)
RMSE.U(Cardiological$Pulse,pred.cm$Fitted)
R2.L(Cardiological$Pulse,pred.cm$Fitted)
R2.U(Cardiological$Pulse,pred.cm$Fitted)
deter.coefficient(Cardiological$Pulse,pred.cm$Fitted)
```

---

cardiologicalv2 *Cardiological data example*

Description

Cardiological interval data example.

Usage

```
data(Cardiological)
```

Format

An object of class `symbolic_tbl` (inherits from `tbl_df`, `tbl`, `data.frame`) with 44 rows and 5 columns.

References

Billard L. and Diday E. (2006).Symbolic data analysis: Conceptual statistics and data mining. Wiley, Chichester.

---

centers.interval *Compute centers of the interval*

Description

Compute centers of the interval

Usage

```
centers.interval(sym.data)
```

Arguments

`sym.data`	Symbolic interval data table.

Value

Centers of teh intervals.

Author(s)

Jorge Arce.

References

Arce J. and Rodriguez O. (2015) 'Principal Curves and Surfaces to Interval Valued Variables'. The 5th Workshop on Symbolic Data Analysis, SDA2015, Orleans, France, November.

Hastie,T. (1984).Principal Curves and Surface. Ph.D Thesis Stanford University.

Hastie,T. & Weingessel,A. (2014). princurve - Fits a Principal Curve in Arbitrary Dimension.R package version 1.1–12 http://cran.r-project.org/web/packages/princurve/index.html.

Hastie,T. & Stuetzle, W. (1989). Principal Curves. Journal of the American Statistical Association, Vol. 84-406, 502–516.

Hastie, T., Tibshirani, R. & Friedman, J. (2008). The Elements of Statistical Learning; Data Mining, Inference and Prediction. Springer, New York.

See Also

sym.interval.pc

`classic.to.sym` *Generate a symbolic data frame*

Description

Generate a symbolic data table from a classic data table.

Usage

```
classic.to.sym(
  x = NULL,
  concept = NULL,
  variables = tidyselect::everything(),
  default.numeric = sym.interval,
  default.categorical = sym.modal,
  ...
)
```

Arguments

`x`	A data.frame.
`concept`	These are the variable that we are going to use a concepts.
`variables`	These are the variables that we want to include in the symbolic data table.
`default.numeric`	function to use for numeric variables
`default.categorical`	function to use for categorical variables
`...`	A vector with names and the type of symbolic data to use, the available types are type_histogram (), type_continuous (), type.set (), type.modal (), by default type_histogram () is used for numeric variables and type_modal () for the categorical variables.

Value

a [tibble][tibble::tibble-package]

References

Bock H-H. and Diday E. (eds.) (2000). Analysis of Symbolic Data. Exploratory methods for extracting statistical information from complex data. Springer, Germany.

---

cor *Generic function for the correlation*

Description

This function compute the symbolic correlation

Usage

```
cor(x, ...)

## Default S3 method:
cor(
  x,
  y = NULL,
  use = "everything",
  method = c("pearson", "kendall", "spearman"),
  ...
)

## S3 method for class 'symbolic_interval'
cor(x, y, method = c("centers", "billard"), ...)

## S3 method for class 'symbolic_tbl'
cor(x, ...)
```

Arguments

`x`	A symbolic variable.
`...`	As in R cor function.
`y`	A symbolic variable.
`use`	An optional character string giving a method for computing covariances in the presence of missing values. This must be (an abbreviation of) one of the strings 'everything', 'all.obs', 'complete.obs', 'na.or.complete', or 'pairwise.complete.obs'.
`method`	The method to be use.

Value

Return a real number in [-1,1].

Author(s)

Oldemar Rodriguez Rojas

References

Billard L. and Diday E. (2006). Symbolic data analysis: Conceptual statistics and data mining. Wiley, Chichester.

Rodriguez, O. (2000). Classification et Modeles Lineaires en Analyse des Donnees Symboliques. Ph.D. Thesis, Paris IX-Dauphine University.

`cov` *Generic function for the covariance*

Description

This function compute the symbolic covariance.

Usage

```
cov(x, ...)

## Default S3 method:
cov(
  x,
  y = NULL,
  use = "everything",
  method = c("pearson", "kendall", "spearman"),
  ...
)

## S3 method for class 'symbolic_interval'
```

```
cov(x, y, method = c("centers", "billard"), na.rm = FALSE, ...)

## S3 method for class 'symbolic_tbl'
cov(x, ...)
```

Arguments

x	First symbolic variables.
...	As in R cov function.
y	Second symbolic variables.
use	an optional character string giving a method for computing covariances in the presence of missing values. This must be (an abbreviation of) one of the strings 'everything', 'all.obs', 'complete.obs', 'na.or.complete', or 'pairwise.complete.obs'.
method	The method to be use.
na.rm	As in R cov function.

Value

Return a real number.

Author(s)

Oldemar Rodriguez Rojas

References

Billard L. and Diday E. (2006). Symbolic data analysis: Conceptual statistics and data mining. Wiley, Chichester.

Rodriguez, O. (2000). Classification et Modeles Lineaires en Analyse des Donnees Symboliques. Ph.D. Thesis, Paris IX-Dauphine University.

---

deter.coefficient *Compute the determination cosfficient*

Description

The determination coefficient represents a goodness-of-fit measure commonly used in regression analysis to capture the adjustment quality of a model.

Usage

```
deter.coefficient(ref, pred)
```

Arguments

ref	Variable that was predicted.
pred	The prediction given by the model.

Value

Return the determination cosfficient.

Author(s)

Oldemar Rodriguez Rojas

References

LIMA-NETO, E.A., DE CARVALHO, F.A.T., (2008). Centre and range method to fitting a linear regression model on symbolic interval data. Computational Statistics and Data Analysis 52, 1500-1515.

LIMA-NETO, E.A., DE CARVALHO, F.A.T., (2010). Constrained linear regression models for symbolic interval-valued variables. Computational Statistics and Data Analysis 54, 333-347.

See Also

sym.glm

Examples

```
data(int_prost_test)
data(int_prost_train)
res.cm <- sym.lm(lpsa ~ ., sym.data = int_prost_train, method = "cm")
pred.cm <- sym.predict(res.cm, int_prost_test)
deter.coefficient(int_prost_test$lpsa, pred.cm$Fitted)
```

---

dist.vect *Compute a distance vector*

Description

Compute a distance vector

Usage

```
dist.vect(vector1, vector2)
```

Arguments

vector1	First vector.
vector2	Second vector.

Value

Eclidean distance between the two vectors.

Author(s)

Jorge Arce

References

Arce J. and Rodriguez O. (2015) 'Principal Curves and Surfaces to Interval Valued Variables'. The 5th Workshop on Symbolic Data Analysis, SDA2015, Orleans, France, November.

Hastie,T. (1984). Principal Curves and Surface. Ph.D. Thesis Stanford University.

Hastie,T. & Weingessel,A. (2014). princurve - Fits a Principal Curve in Arbitrary Dimension.R package version 1.1–12 http://cran.r-project.org/web/packages/princurve/index.html.

Hastie,T. & Stuetzle, W. (1989). Principal Curves. Journal of the American Statistical Association, Vol. 84-406, 502–516.

Hastie, T., Tibshirani, R. & Friedman, J. (2008). The Elements of Statistical Learning; Data Mining, Inference and Prediction. Springer, New York.

See Also

sym.interval.pc

---

dist.vect.matrix *Compute the distance vector matrix*

Description

Compute the distance vector matrix.

Usage

```
dist.vect.matrix(vector, Matrix)
```

Arguments

`vector`	An n dimensional vector.
`Matrix`	An n x n matrix.

Value

The distance.

Author(s)

Jorge Arce.

References

Arce J. and Rodriguez O. (2015) 'Principal Curves and Surfaces to Interval Valued Variables'. The 5th Workshop on Symbolic Data Analysis, SDA2015, Orleans, France, November.

Hastie,T. (1984). Principal Curves and Surface. Ph.D Thesis Stanford University.

Hastie,T. & Weingessel,A. (2014). princurve - Fits a Principal Curve in Arbitrary Dimension.R package version 1.1–12 http://cran.r-project.org/web/packages/princurve/index.html.

Hastie,T. & Stuetzle, W. (1989). Principal Curves. Journal of the American Statistical Association, Vol. 84-406, 502–516.

Hastie, T., Tibshirani, R. & Friedman, J. (2008). The Elements of Statistical Learning; Data Mining, Inference and Prediction. Springer, New York.

See Also

sym.interval.pc

---

ex1_db2so — *Data example to generate symbolic objets*

Description

This is a small data example to generate symbolic objets.

Usage

```
data(ex1_db2so)
```

Format

An object of class `data.frame` with 19 rows and 5 columns.

References

Bock H-H. and Diday E. (eds.) (2000). Analysis of Symbolic Data. Exploratory methods for extracting statistical information from complex data. Springer, Germany.

Examples

```
data(ex1_db2so)
ex1 <- ex1_db2so
result <- classic.to.sym(
  x = ex1_db2so,
  concept = c(state, sex),
  variables = c(county, group, age),
  county = mean(county),
  age_hist = sym.histogram(age, breaks = pretty(ex1_db2so$age, 5))
)
result
```

`example1` *Data Example 1*

Description

This a symbolic data table with variables of continuos, interval, histogram and set types.

Usage

```
data(example1)
```

Format

The labels $C means that follows a continuous variable, $I means an interval variable, $H means a histogram variables and $S means set variable. In the first row each labels should be follow of a name to variable and to the case of histogram a set variables types the names of the modalities (categories). In data rows for continuous variables we have just one value, for interval variables we have the minimum and the maximum of the interval, for histogram variables we have the number of modalities and then the probability of each modality and for set variables we have the cardinality of the set and next the elements of the set.
The format is the *.csv file is:
$C F1 $I F2 F2 $M F3 M1 M2 M3 $S F4 e a 2 3 g b 1 4 i k c d
Case1 $C 2.8 $I 1 2 $M 3 0.1 0.7 0.2 $S 12 1 0 0 0 1 0 0 0 1 1 0 0
Case2 $C 1.4 $I 3 9 $M 3 0.6 0.3 0.1 $S 12 0 1 0 0 0 1 0 0 0 0 1 1
Case3 $C 3.2 $I -1 4 $M 3 0.2 0.2 0.6 $S 12 0 0 1 0 0 1 1 0 0 0 1 0
Case4 $C -2.1 $I 0 2 $M 3 0.9 0.0 0.1 $S 12 0 1 0 1 0 0 0 1 0 0 1 0
Case5 $C -3.0 $I -4 -2 $M 3 0.6 0.0 0.4 $S 12 1 0 0 0 1 0 0 0 1 1 0 0
The internal format is:
$N
[1] 5
$M
[1] 4
$sym.obj.names
[1] 'Case1' 'Case2' 'Case3' 'Case4' 'Case5'
$sym.var.names
[1] 'F1' 'F2' 'F3' 'F4'
$sym.var.types [1] '$C' '$I' '$H' '$S'
$sym.var.length
[1] 1 2 3 4
$sym.var.starts
[1] 2 4 8 13
$meta

$C F1 $I F2 F2 $M F3 M1 M2 M3 $S F4 e a 2 3 g b 1 4 i k c d Case1 $C 2.8 $I 1 2 $M 3 0.1 0.7 0.2 $S 12 1 0 0 0 1 0 0 0 1 1 0 0 Case2 $C 1.4 $I 3 9 $M 3 0.6 0.3 0.1 $S 12 0 1 0 0 0 1 0 0 0 0 1 1 Case3 $C 3.2 $I -1 4 $M 3 0.2 0.2 0.6 $S 12 0 0 1 0 0 1 1 0 0 0 1 0 Case4 $C -2.1 $I 0 2 $M 3 0.9 0.0 0.1 $S 12 0 1 0 1 0 0 0 1 0 0 1 0 Case5 $C -3.0 $I -4 -2 $M 3 0.6 0.0 0.4 $S 12 1 0 0 0 1 0 0 0 1

1 0 0 $data
F1 F2 F2.1 M1 M2 M3 e a 2 3 g b 1 4 i k c d Case1 2.8 1 2 0.1 0.7 0.2 1 0 0 0 1 0 0 0 1 1 0 0 Case2 1.4 3 9 0.6 0.3 0.1 0 1 0 0 0 1 0 0 0 0 1 1 Case3 3.2 -1 4 0.2 0.2 0.6 0 0 1 0 0 1 1 0 0 0 1 0 Case4 -2.1 0 2 0.9 0.0 0.1 0 1 0 1 0 0 0 1 0 0 1 0 Case5 -3.0 -4 -2 0.6 0.0 0.4 1 0 0 0 1 0 0 0 1 1 0 0

References

Bock H-H. and Diday E. (eds.) (2000). Analysis of Symbolic Data. Exploratory methods for extracting statistical information from complex data. Springer, Germany.

Examples

```
data(example1)
example1
```

---

example2 *Data Example 2*

Description

This a symbolic data table with variables of continuos, interval, histogram and set types.

Usage

```
data(example2)
```

Format

$C F1 $I F2 F2 $M F3 M1 M2 M3 $C F4 $S F5 e a 2 3 g b 1 4 i k c d
Case1 $C 2.8 $I 1 2 $M 3 0.1 0.7 0.2 $C 6.0 $S 12 1 0 0 0 1 0 0 0 1 1 0 0
Case2 $C 1.4 $I 3 9 $M 3 0.6 0.3 0.1 $C 8.0 $S 12 0 1 0 0 0 1 0 0 0 0 1 1
Case3 $C 3.2 $I -1 4 $M 3 0.2 0.2 0.6 $C -7.0 $S 12 0 0 1 0 0 1 1 0 0 0 1 0
Case4 $C -2.1 $I 0 2 $M 3 0.9 0.0 0.1 $C 0.0 $S 12 0 1 0 1 0 0 0 1 0 0 1 0
Case5 $C -3.0 $I -4 -2 $M 3 0.6 0.0 0.4 $C -9.5 $S 12 1 0 0 0 1 0 0 0 1 1 0 0

Examples

```
data(example2)
example2
```

`example3` *Data Example 3*

Description

This a symbolic data table with variables of continuos, interval, histogram and set types.

Usage

```
data(example3)
```

Format

$C F1 $I F2 F2 $M F3 M1 M2 M3 $C F4 $S F5 e a 2 3 g b 1 4 i k c d $I F6 F6 $I F7 F7 Case1 $C 2.8 $I 1 2 $M 3 0.1 0.7 0.2 $C 6.0 $S 12 1 0 0 0 1 0 0 0 1 1 0 0 $I 0.00 90.00 $I 9 24 Case2 $C 1.4 $I 3 9 $M 3 0.6 0.3 0.1 $C 8.0 $S 12 0 1 0 0 0 1 0 0 0 0 1 1 $I -90.00 98.00 $I -9 9 Case3 $C 3.2 $I -1 4 $M 3 0.2 0.2 0.6 $C -7.0 $S 12 0 0 1 0 0 1 1 0 0 0 1 0 $I 65.00 90.00 $I 65 70 Case4 $C -2.1 $I 0 2 $M 3 0.9 0.0 0.1 $C 0.0 $S 12 0 1 0 1 0 0 0 1 0 0 1 0 $I 45.00 89.00 $I 25 67 Case5 $C -3.0 $I -4 -2 $M 3 0.6 0.0 0.4 $C -9.5 $S 12 1 0 0 0 1 0 0 0 1 1 0 0 $I 20.00 40.00 $I 9 40 Case6 $C 0.1 $I 10 21 $M 3 0.0 0.7 0.3 $C -1.0 $S 12 1 0 0 0 0 0 1 0 1 0 0 0 $I 5.00 8.00 $I 5 8 Case7 $C 9.0 $I 4 21 $M 3 0.2 0.2 0.6 $C 0.5 $S 12 1 1 1 0 0 0 0 0 0 0 0 0 $I 3.14 6.76 $I 4 6

Examples

```
data(example3)
example3
```

`example4` *Data Example 4*

Description

data(example4) example4

Usage

```
data(example4)
```

Format

$C 2.8 $I 1 2 $M 3 0.1 0.7 0.2 $C 6 $S F4 e a 2 3 g b 1 4 i k c d $I 0 90 Case2 $C 1.4 $I 3 9 $M 3 0.6 0.3 0.1 $C 8.0 $S 12 1 0 0 0 1 0 0 0 1 1 0 0 $I -90.00 98.00 Case3 $C 3.2 $I -1 4 $M 3 0.2 0.2 0.6 $C -7.0 $S 12 0 1 0 0 0 1 0 0 0 0 1 1 $I 65.00 90.00 Case4 $C -2.1 $I 0 2 $M 3 0.9 0.0 0.1 $C 0.0 $S 12 0 0 1 0 0 1 1 0 0 0 1 0 $I 45.00 89.00 Case5 $C -3.0 $I -4 -2 $M 3 0.6 0.0 0.4 $C -9.5 $S 12 0 1 0 1 0 0 0 1 0 0 1 0 $I 90.00 990.00 Case6 $C 0.1 $I 10 21 $M 3 0.0 0.7 0.3 $C -1.0 $S 12 1 0 0 0 1 0 0 0 1 1 0 0 $I 5.00 8.00 Case7 $C 9.0 $I 4 21 $M 3 0.2 0.2 0.6 $C 0.5 $S 12 1 1 0 0 0 0 1 0 0 0 0 1 $I 3.14 6.76

Examples

```
data(example4)
example4
```

---

example5 *Data Example 5*

Description

This a symbolic data matrix wint continuos, interval, histograma a set data types.

Usage

```
data(example5)
```

Format

$H F0 M01 M02 $C F1 $I F2 F2 $H F3 M1 M2 M3 $S F4 E1 E2 E3 E4
Case1 $H 2 0.1 0.9 $C 2.8 $I 1 2 $H 3 0.1 0.7 0.2 $S 4 e g k i
Case2 $H 2 0.7 0.3 $C 1.4 $I 3 9 $H 3 0.6 0.3 0.1 $S 4 a b c d
Case3 $H 2 0.0 1.0 $C 3.2 $I -1 4 $H 3 0.2 0.2 0.6 $S 4 2 1 b c
Case4 $H 2 0.2 0.8 $C -2.1 $I 0 2 $H 3 0.9 0.0 0.1 $S 4 3 4 c a
Case5 $H 2 0.6 0.4 $C -3.0 $I -4 -2 $H 3 0.6 0.0 0.4 $S 4 e i g k

Examples

```
data(example5)
example5
```

---

example6 *Data Example 6*

Description

This a symbolic data matrix wint continuos, interval, histograma a set data types.

Usage

```
data(example6)
```

Format

$C F1 $M F2 M1 M2 M3 M4 M5 $I F3 F3 $M F4 M1 M2 M3 $C F5 $S F4 e a 2 3 g b 1 4 i k c d
Case1 $C 2.8 $M 5 0.1 0.1 0.1 0.1 0.6 $I 1 2 $M 3 0.1 0.7 0.2 $C 6.0 $S 12 1 0 0 0 1 0 0 0 1 1 0 0
Case2 $C 1.4 $M 5 0.1 0.1 0.1 0.1 0.6 $I 3 9 $M 3 0.6 0.3 0.1 $C 8.0 $S 12 0 1 0 0 0 1 0 0 0 0 1 1
Case3 $C 3.2 $M 5 0.1 0.1 0.1 0.1 0.6 $I -1 4 $M 3 0.2 0.2 0.6 $C -7.0 $S 12 0 0 1 0 0 1 1 0 0 0 1 0
Case4 $C -2.1 $M 5 0.1 0.1 0.1 0.1 0.6 $I 0 2 $M 3 0.9 0.0 0.1 $C 0.0 $S 12 0 1 0 1 0 0 0 1 0 0 1 0
Case5 $C -3.0 $M 5 0.1 0.1 0.1 0.1 0.6 $I -4 -2 $M 3 0.6 0.0 0.4 $C -9.5 $S 12 1 0 0 0 1 0 0 0 1 1 0 0

Examples

```
data(example6)
example6
```

`example7` *Data Example 7*

Description

This a symbolic data matrix wint continuos, interval, histograma a set data types.

Usage

```
data(example6)
```

Format

$C F1 $H F2 M1 M2 M3 M4 M5 $I F3 F3 $H F4 M1 M2 M3 $C F5
Case1 $C 2.8 $H 5 0.1 0.2 0.3 0.4 0.0 $I 1 2 $H 3 0.1 0.7 0.2 $C 6.0
Case2 $C 1.4 $H 5 0.2 0.1 0.5 0.1 0.2 $I 3 9 $H 3 0.6 0.3 0.1 $C 8.0
Case3 $C 3.2 $H 5 0.1 0.1 0.2 0.1 0.5 $I -1 4 $H 3 0.2 0.2 0.6 $C -7.0
Case4 $C -2.1 $H 5 0.4 0.1 0.1 0.1 0.3 $I 0 2 $H 3 0.9 0.0 0.1 $C 0.0
Case5 $C -3.0 $H 5 0.6 0.1 0.1 0.1 0.1 $I -4 -2 $H 3 0.6 0.0 0.4 $C -9.5

Examples

```
data(example7)
example7
```

`ex_cfa1` *Correspondence Analysis Example*

Description

Correspondence Analysis for Symbolic MultiValued Variables example.

Usage

```
data(ex_cfa1)
```

Format

An object of class `symbolic_tbl` (inherits from `tbl_df`, `tbl`, `data.frame`) with 4 rows and 4 columns.

References

Rodriguez, O. (2011). Correspondence Analysis for Symbolic MultiValued Variables. Workshop in Symbolic Data Analysis Namur, Belgium

`ex_cfa2` *Correspondence Analysis Example*

Description

Correspondence Analysis for Symbolic MultiValued Variables example.

Usage

```
data(ex_cfa2)
```

Format

An object of class `symbolic_tbl` (inherits from `tbl_df`, `tbl`, `data.frame`) with 6 rows and 5 columns.

References

Rodriguez, O. (2011). Correspondence Analysis for Symbolic MultiValued Variables. Workshop in Symbolic Data Analysis Namur, Belgium

ex_mcfa1 *Multiple Correspondence Analysis Example*

Description

example for the sym.mcfa function.

example for the sym.mcfa function.

Usage

```
data(ex_mcfa1)

ex_mcfa1
```

Format

An object of class `data.frame` with 130 rows and 5 columns.

An object of class `data.frame` with 130 rows and 5 columns.

Examples

```
data("ex_mcfa1")
sym.table <- classic.to.sym(ex_mcfa1,
                            concept = suspect,
                            hair = sym.set(hair),
                            eyes = sym.set(eyes),
                            region = sym.set(region))

res <- sym.mcfa(sym.table, c(1,2))
mcfa.scatterplot(res[,1], res[,2], sym.data = sym.table, pos.var = c(1,2))

data("ex_mcfa1")
sym.table <- classic.to.sym(
  x = ex_mcfa1,
  concept = "suspect",
  variables = c(hair, eyes, region),
  hair = sym.set(hair),
  eyes = sym.set(eyes),
  region = sym.set(region)
)
sym.table
```

ex_mcfa2 *Multiple Correspondence Analysis Example*

Description

example for the sym.mcfa function.

Usage

```
data(ex_mcfa2)
```

Format

An object of class `data.frame` with 130 rows and 7 columns.

Examples

```
data("ex_mcfa2")

ex <- classic.to.sym(ex_mcfa2,
                    concept = employee_id,
                   variables = c(employee_id, salary, region, evaluation, years_worked),
                     salary = sym.set(salary),
                     region = sym.set(region),
                     evaluation = sym.set(evaluation),
                     years_worked = sym.set(years_worked))

res <- sym.mcfa(ex, c(1,2,3,4))
mcfa.scatterplot(res[,1], res[,2], sym.data = ex, pos.var = c(1,2,3,4))
```

facedata *Face Data Example*

Description

Symbolic data matrix with all the variables of interval type.

Usage

```
data('facedata')
```

Format

$I;AD;AD;$I;BC;BC;.........

HUS1;$I;168.86;172.84;$I;58.55;63.39;.........
HUS2;$I;169.85;175.03;$I;60.21;64.38;.........
HUS3;$I;168.76;175.15;$I;61.4;63.51;.........
INC1;$I;155.26;160.45;$I;53.15;60.21;.........
INC2;$I;156.26;161.31;$I;51.09;60.07;.........
INC3;$I;154.47;160.31;$I;55.08;59.03;.........
ISA1;$I;164;168;$I;55.01;60.03;.........
ISA2;$I;163;170;$I;54.04;59;.........
ISA3;$I;164.01;169.01;$I;55;59.01;.........
JPL1;$I;167.11;171.19;$I;61.03;65.01;.........
JPL2;$I;169.14;173.18;$I;60.07;65.07;.........
JPL3;$I;169.03;170.11;$I;59.01;65.01;.........
KHA1;$I;149.34;155.54;$I;54.15;59.14;.........
KHA2;$I;149.34;155.32;$I;52.04;58.22;.........
KHA3;$I;150.33;157.26;$I;52.09;60.21;.........
LOT1;$I;152.64;157.62;$I;51.35;56.22;.........
LOT2;$I;154.64;157.62;$I;52.24;56.32;.........
LOT3;$I;154.83;157.81;$I;50.36;55.23;.........
PHI1;$I;163.08;167.07;$I;66.03;68.07;.........
PHI2;$I;164;168.03;$I;65.03;68.12;.........
PHI3;$I;161.01;167;$I;64.07;69.01;.........
ROM1;$I;167.15;171.24;$I;64.07;68.07;.........
ROM2;$I;168.15;172.14;$I;63.13;68.07;.........
ROM3;$I;167.11;171.19;$I;63.13;68.03;.........

References

Billard L. and Diday E. (2006). Symbolic data analysis: Conceptual statistics and data mining. Wiley, Chichester.

Examples

```
## Not run:
data(facedata)
res.vertex.ps <- sym.interval.pc(facedata,'vertex',150,FALSE,FALSE,TRUE)
class(res.vertex.ps$sym.prin.curve) <- c('sym.data.table')
sym.scatterplot(res.vertex.ps$sym.prin.curve[,1], res.vertex.ps$sym.prin.curve[,2],
                labels=TRUE,col='red',main='PSC Face Data')

## End(Not run)
```

format.symbolic_histogram

Symbolic modal conversion functions to and from Character

Description

Symbolic modal conversion functions to and from Character

Usage

```
## S3 method for class 'symbolic_histogram'
format(x, ...)
```

Arguments

x	An object to be converted
...	Further arguments to be passed from or to other methods.

format.symbolic_interval

Symbolic interval conversion functions to and from Character

Description

Symbolic interval conversion functions to and from Character

Usage

```
## S3 method for class 'symbolic_interval'
format(x, ...)
```

Arguments

x	An object to be converted
...	Further arguments to be passed from or to other methods.

format.symbolic_modal *Symbolic modal conversion functions to and from Character*

Description

Symbolic modal conversion functions to and from Character

Usage

```
## S3 method for class 'symbolic_modal'
format(x, ...)
```

Arguments

x	An object to be converted
...	Further arguments to be passed from or to other methods.

format.symbolic_set *Symbolic set conversion functions to and from Character*

Description

Symbolic set conversion functions to and from Character

Usage

```
## S3 method for class 'symbolic_set'
format(x, ...)
```

Arguments

x	An object to be converted
...	Further arguments to be passed from or to other methods.

get.limits.PCA *Projections onto PCA*

Description

Calculate the interval projection onto the principal components

Usage

```
get.limits.PCA(sym.data, matrix.stan, min.stan, max.stan, svd, nn, mm)
```

Arguments

sym.data	An interval matrix
matrix.stan	A standardized matrix
min.stan	A matrix of minimum values standardized for each interval
max.stan	A matrix of maximum values standardized for each interval
svd	An eigen vectors matrix
nn	Number of concepts
mm	Number of variables

Value

Concept Projections onto the principal components and correlation circle

get.limits.PCA.indivduals *Projections onto PCA*

Description

Calculate the interval projection onto the principal components

Usage

```
get.limits.PCA.indivduals(
  sym.data,
  matrix.stan,
  min.stan,
  max.stan,
  svd,
  nn,
  mm
)
```

Arguments

`sym.data`	An interval matrix
`matrix.stan`	A standardized matrix
`min.stan`	A matrix of minimum values standardized for each interval
`max.stan`	A matrix of maximum values standardized for each interval
`svd`	An eigen vectors matrix
`nn`	Number of concepts
`mm`	Number of variables

Value

Concept Projections onto the principal components

`get_cats` *Extract categories*

Description

Extract categories

Usage

```
get_cats(x, ...)
```

Arguments

`x`	An object to be converted
`...`	Further arguments to be passed from or to other methods.

`get_props` *Extract prop*

Description

Extract prop

Usage

```
get_props(x, ...)
```

Arguments

`x`	An object to be converted
`...`	Further arguments to be passed from or to other methods.

`hardwoodBrito` *Hard Wood Data Example*

Description

Symbolic Histogram matrix.

Usage

```
data('hardwoodBrito')
```

Format

An object of class `symbolic_tbl` (inherits from `symbolic_tbl`, `symbolic_tbl`, `symbolic_tbl`, `tbl_df`, `tbl`, `data.frame`) with 5 rows and 4 columns.

References

Brito P. and Dias S. (2022). Analysis of Distributional Data. CRC Press, United States of America.

Examples

```
## Not run:
data(hardwoodBrito)
hardwoodBrito

## End(Not run)
```

`HistRSDAToEcdf` *HistRSDAToEcdf*

Description

HistRSDAToEcdf

Usage

```
HistRSDAToEcdf(h)
```

Arguments

`h`	A matrix of histograms

Value

Transformation in Ecdf object

Author(s)

Jorge Arce Garro

Examples

```
## Not run:
 data("hardwoodBrito")
 Hardwood.histogram<-hardwoodBrito
 Hardwood.cols<-colnames(Hardwood.histogram)
 Hardwood.names<-row.names(Hardwood.histogram)
 M<-length(Hardwood.cols)
 N<-length(Hardwood.names)
 BIN.Matrix<-matrix(rep(3,N*M),nrow = N)
 pca.hist<-sym.histogram.pca(Hardwood.histogram,BIN.Matrix)
 Hardwood.quantiles.PCA.2<-quantiles.RSDA.KS(pca.hist$sym.hist.matrix.PCA,100)
 h<-Hardwood.quantiles.PCA.2[[1]][[1]]
 HistRSDAToEcdf(h)

## End(Not run)
```

`interval.centers` *calcula centros*

Description

calcula centros

Usage

```
interval.centers(x)
```

Arguments

x	tabla simbolica todos intervalos

`interval.histogram.plot` *Histogram plot for an interval variable*

Description

Histogram plot for an interval variable

Usage

```
interval.histogram.plot(x, n.bins, ...)
```

Arguments

`x`	An symbolic data table.
`n.bins`	Numbers of breaks of the histogram.
`...`	Arguments to be passed to the barplot method.

Value

A list with componets : frequency and histogram

Examples

```
data(oils)
res <- interval.histogram.plot(x = oils[, 3], n.bins = 3)
res
```

`interval.large` *Calculate the large of each interval*

Description

Calculate the large of each interval

Usage

```
interval.large(x)
```

Arguments

`x`	An interval matrix

Value

A matrix with the large of each interval.

Examples

```
## Not run:
data(oils)
interval.large(oils)

## End(Not run)
```

`interval.length` *Lenght for interval*

Description

Calculate the large of each interval

Usage

```
interval.length(x)
```

Arguments

`x`	An interval matrix

Value

A matrix with the length of each interval.

Examples

```
## Not run:
data(oils)
interval.length(oils)

## End(Not run)
```

`interval.max` *calcula maximos*

Description

calcula maximos

Usage

```
interval.max(x)
```

Arguments

`x`	tabla simbolica todos intervalos

interval.min *calcula minimos*

Description

calcula minimos

Usage

```
interval.min(x)
```

Arguments

`x`	tabla simbolica todos intervalos

interval.ranges *calcula rangos*

Description

calcula rangos

Usage

```
interval.ranges(x)
```

Arguments

`x`	tabla simbolica todos intervalos

int_prost_test *Linear regression model data example.*

Description

Linear regression model interval-valued data example.

Usage

```
data(int_prost_test)
```

Format

An object of class `symbolic_tbl` (inherits from `tbl_df`, `tbl`, `data.frame`) with 30 rows and 9 columns.

References

HASTIE, T., TIBSHIRANI, R. and FRIEDMAN, J. (2008). The Elements of Statistical Learning: Data Mining, Inference and Prediction. New York: Springer.

---

int_prost_train *Linear regression model data example.*

Description

Linear regression model interval-valued data example.

Usage

```
data(int_prost_train)
```

Format

An object of class `symbolic_tbl` (inherits from `tbl_df`, `tbl`, `data.frame`) with 67 rows and 9 columns.

References

HASTIE, T., TIBSHIRANI, R. and FRIEDMAN, J. (2008). The Elements of Statistical Learning: Data Mining, Inference and Prediction. New York: Springer.

---

is.sym.histogram *Symbolic histogram*

Description

Symbolic histogram

Usage

```
is.sym.histogram(x)
```

Arguments

`x`	an object to be tested

Value

returns TRUE if its argument's value is a symbolic_histogram and FALSE otherwise.

Examples

```
x <- sym.histogram(iris$Sepal.Length)
is.sym.histogram(x)
```

is.sym.interval *Symbolic interval*

Description

Symbolic interval

Usage

```
is.sym.interval(x)
```

Arguments

x	an object to be tested

Value

returns TRUE if its argument's value is a symbolic_vector and FALSE otherwise.

Examples

```
x <- sym.interval(1:10)
is.sym.interval(x)
is.sym.interval("d")
```

is.sym.modal *Symbolic modal*

Description

Symbolic modal

Usage

```
is.sym.modal(x)
```

Arguments

x	an object to be tested

Value

returns TRUE if its argument's value is a symbolic_modal and FALSE otherwise.

Examples

```
x <- sym.modal(factor(c("a", "b", "b", "l")))
is.sym.modal(x)
```

`is.sym.set` *Symbolic set*

Description

Symbolic set

Usage

```
is.sym.set(x)
```

Arguments

x	an object to be tested

Value

returns TRUE if its argument's value is a symbolic_set and FALSE otherwise.

Examples

```
x <- sym.set(factor(c("a", "b", "b", "l")))
is.sym.set(x)
```

`lynne1` *Symbolic interval data example.*

Description

Symbolic data matrix with all the variables of interval type.

Usage

```
data(lynne1)
```

Format

An object of class `tbl_df` (inherits from `tbl`, `data.frame`) with 10 rows and 4 columns.

References

Billard L. and Diday E. (2006). Symbolic data analysis: Conceptual statistics and data mining. Wiley, Chichester.

Examples

```
data(lynne1)
lynne1
```

mcfa.scatterplot *Plot Interval Scatterplot*

Description

Plot Interval Scatterplot

Usage

```
mcfa.scatterplot(x, y, sym.data, pos.var)
```

Arguments

x	symbolic table with only one column.
y	symbolic table with only one column.
sym.data	original symbolic table.
pos.var	column number of the variables to be plotted.

Examples

```
data("ex_mcfa1")
sym.table <- classic.to.sym(ex_mcfa1,
  concept = suspect,
  hair = sym.set(hair),
  eyes = sym.set(eyes),
  region = sym.set(region)
)

res <- sym.mcfa(sym.table, c(1, 2))
mcfa.scatterplot(res[, 2], res[, 3], sym.data = sym.table, pos.var = c(1, 2))
```

mean.symbolic_interval *Symbolic mean for intervals*

Description

This function compute the symbolic mean for intervals

Usage

```
## S3 method for class 'symbolic_interval'
mean(x, method = c("centers", "interval"), trim = 0, na.rm = F, ...)

## S3 method for class 'symbolic_tbl'
mean(x, ...)
```

Arguments

`x`	A symbolic interval.
`method`	The method to be use.
`trim`	As in R mean function.
`na.rm`	As in R mean function.
`...`	As in R mean function.

Author(s)

Oldemar Rodriguez Rojas

References

Billard L. and Diday E. (2006). Symbolic data analysis: Conceptual statistics and data mining. Wiley, Chichester.

Rodriguez, O. (2000). Classification et Modeles Lineaires en Analyse des Donnees Symboliques. Ph.D. Thesis, Paris IX-Dauphine University.

---

median.symbolic_interval
Symbolic Median

Description

This function compute the median for symbolic intervals.

Usage

```
## S3 method for class 'symbolic_interval'
median(x, na.rm = FALSE, method = c("centers", "interval"), ...)

## S3 method for class 'symbolic_tbl'
median(x, ...)
```

Arguments

`x`	A symbolic interval.
`na.rm`	As in R median function.
`method`	The method to be use.
`...`	As in R median function.

Author(s)

Oldemar Rodriguez Rojas

References

Billard L. and Diday E. (2006). Symbolic data analysis: Conceptual statistics and data mining. Wiley, Chichester.

Rodriguez, O. (2000). Classification et Modeles Lineaires en Analyse des Donnees Symboliques. Ph.D. Thesis, Paris IX-Dauphine University.

`method_summary` *Summary method to CM and CRM regression model*

Description

Summary method to CM and CRM regression model

Usage

```
method_summary(ref, pred)
```

Arguments

`ref`	Real values
`pred`	Predicted values

`min.symbolic_interval` *Maxima and Minima*

Description

Maxima and Minima

Usage

```
## S3 method for class 'symbolic_interval'
min(x, ...)

## S3 method for class 'symbolic_interval'
max(x, ...)

## S3 method for class 'symbolic_interval'
x$name = c("min", "max", "mean", "median")
```

Arguments

`x`	symbolic interval vector
`...`	further arguments passed to or from other methods.
`name`	...

Value

a new symbolic interval with the minimum of the minima and the minimum of the maxima

---

`neighbors.vertex` *Compute neighbors vertex*

Description

Compute neighbors vertex

Usage

```
neighbors.vertex(vertex, Matrix, num.neig)
```

Arguments

`vertex`	Vertes of the hipercube
`Matrix`	Interval Data Matrix.
`num.neig`	Number of vertices.

Author(s)

Jorge Arce

References

Arce J. and Rodriguez O. (2015) 'Principal Curves and Surfaces to Interval Valued Variables'. The 5th Workshop on Symbolic Data Analysis, SDA2015, Orleans, France, November.

Hastie,T. (1984). Principal Curves and Surface. Ph.D Thesis Stanford University.

Hastie,T. & Weingessel,A. (2014). princurve - Fits a Principal Curve in Arbitrary Dimension.R package version 1.1–12 http://cran.r-project.org/web/packages/princurve/index.html.

Hastie,T. & Stuetzle, W. (1989). Principal Curves. Journal of the American Statistical Association, Vol. 84-406, 502–516.

Hastie, T., Tibshirani, R. & Friedman, J. (2008). The Elements of Statistical Learning; Data Mining, Inference and Prediction. Springer, New York.

See Also

sym.interval.pc

`norm.vect` *Compute the norm of a vector.*

Description

Compute the norm of a vector.

Usage

```
norm.vect(vector1)
```

Arguments

`vector1` An n dimensional vector.

Value

The L2 norm of the vector.

Author(s)

Jorge Arce

References

Arce J. and Rodriguez O. (2015) 'Principal Curves and Surfaces to Interval Valued Variables'. The 5th Workshop on Symbolic Data Analysis, SDA2015, Orleans, France, November.

Hastie,T. (1984). Principal Curves and Surface. Ph.D Thesis Stanford University.

Hastie,T. & Weingessel,A. (2014). princurve - Fits a Principal Curve in Arbitrary Dimension.R package version 1.1–12 http://cran.r-project.org/web/packages/princurve/index.html.

Hastie,T. & Stuetzle, W. (1989). Principal Curves. Journal of the American Statistical Association, Vol. 84-406, 502–516.

Hastie, T., Tibshirani, R. & Friedman, J. (2008). The Elements of Statistical Learning; Data Mining, Inference and Prediction. Springer, New York.

See Also

sym.interval.pc

`oils` *Ichino Oils example data.*

Description

Symbolic data matrix with all the variables of interval type.

Usage

```
data(oils)
```

Format

$I GRA GRA $I FRE FRE $I IOD IOD $I SAP SAP
L $I 0.930 0.935 $I -27 -18 $I 170 204 $I 118 196
P $I 0.930 0.937 $I -5 -4 $I 192 208 $I 188 197
Co $I 0.916 0.918 $I -6 -1 $I 99 113 $I 189 198
S $I 0.920 0.926 $I -6 -4 $I 104 116 $I 187 193
Ca $I 0.916 0.917 $I -25 -15 $I 80 82 $I 189 193
O $I 0.914 0.919 $I 0 6 $I 79 90 $I 187 196
B $I 0.860 0.870 $I 30 38 $I 40 48 $I 190 199
H $I 0.858 0.864 $I 22 32 $I 53 77 $I 190 202

References

Cazes P., Chouakria A., Diday E. et Schektman Y. (1997). Extension de l'analyse en composantes principales a des donnees de type intervalle, Rev. Statistique Appliquee, Vol. XLV Num. 3 pag. 5-24, France.

Examples

```
data(oils)
oils
```

`pca.supplementary.vertex.fun.j.new` *Calculate the distance*

Description

Calculate the distance

Usage

```
pca.supplementary.vertex.fun.j.new(
  x,
  N,
  M,
  sym.var.names,
  sym.data.vertex.matrix,
  tot.individuals
)
```

Arguments

`x`	A Matrix
`N`	Number of concepts
`M`	Number of variables
`sym.var.names`	Names of concepts
`sym.data.vertex.matrix`	Vertex Matrix
`tot.individuals`	Number of individuals

Value

Distance

`Percentil.Arrow.plot` *Percentil.Arrow.plot*

Description

Percentil.Arrow.plot

Usage

```
Percentil.Arrow.plot(
  quantiles.sym,
  concept.names,
  var.names,
  Title,
  axes.x.label,
  axes.y.label,
  label.name
)
```

Arguments

`quantiles.sym`	Matrix of Quantiles
`concept.names`	Concept Names
`var.names`	Variables to plot the arrows
`Title`	Plot title
`axes.x.label`	Label of axis X
`axes.y.label`	Label of axis Y
`label.name`	Label

Value

Arrow Plot

Author(s)

Jorge Arce Garro

Examples

```
## Not run:
data("hardwoodBrito")
Hardwood.histogram<-hardwoodBrito
Hardwood.cols<-colnames(Hardwood.histogram)
Hardwood.names<-row.names(Hardwood.histogram)
 M<-length(Hardwood.cols)
 N<-length(Hardwood.names)
 BIN.Matrix<-matrix(rep(3,N*M),nrow = N)
pca.hist<-sym.histogram.pca(Hardwood.histogram,BIN.Matrix)
 M<-length(Hardwood.cols)
 N<-length(Hardwood.names)
 BIN.Matrix<-matrix(rep(3,N*M),nrow = N)
label.name<-"Hard Wood"
Title<-"First Principal Plane"
axes.x.label<- "First Principal Component (84.83%)"
axes.y.label<- "Second Principal Component (9.70%)"
concept.names<-c("ACER")
var.names<-c("PC.1","PC.2")
quantile.ACER.plot<-Percentil.Arrow.plot(Hardwood.quantiles.PCA,
                                         concept.names,
                                         var.names,
                                         Title,
                                         axes.x.label,
                                         axes.y.label,
                                         label.name
                                         )
quantile.ACER.plot

## End(Not run)
```

plot.symbolic_tbl *Function for plotting a symbolic object*

Description

Function for plotting a symbolic object

Usage

```
## S3 method for class 'symbolic_tbl'
plot(
  x,
  col = NA,
  matrix.form = NA,
  border = FALSE,
  size = 1,
  title = TRUE,
  show.type = FALSE,
  font.size = 1,
  reduce = FALSE,
  hist.angle.x = 60,
  ...
)
```

Arguments

x	The symbolic object.
col	A specification for the default plotting color.
matrix.form	A vector of the form c(num.rows,num.columns).
border	A logical value indicating whether border should be plotted.
size	The magnification to be used for each graphic.
title	A logical value indicating whether title should be plotted.
show.type	A logical value indicating whether type should be plotted.
font.size	The font size of graphics.
reduce	A logical value indicating whether values different from zero should be plotted in modal and set graphics.
hist.angle.x	The angle of labels in y axis. Only for histogram plot
...	Arguments to be passed to methods.

Value

A plot of the symbolic data table.

Author(s)

Andres Navarro

Examples

```
## Not run:
data(oils)
plot(oils)
plot(oils, border = T, size = 1.3)

## End(Not run)
```

---

plot.sym_umap *Plot UMAP for symbolic data tables*

Description

Plot UMAP for symbolic data tables

Usage

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

Arguments

x	sym_umap object
...	params for plot

---

quantiles.RSDA *quantiles.RSDA*

Description

quantiles.RSDA

Usage

```
quantiles.RSDA(histogram.matrix, num.quantiles)
```

Arguments

histogram.matrix	A matrix of histograms
num.quantiles	Number of quantiles

Value

Quantiles of a Histogram Matrix

Author(s)

Jorge Arce Garro

Examples

```
## Not run:
data("hardwoodBrito")
Hardwood.histogram<-hardwoodBrito
Hardwood.cols<-colnames(Hardwood.histogram)
Hardwood.names<-row.names(Hardwood.histogram)
 M<-length(Hardwood.cols)
 N<-length(Hardwood.names)
 BIN.Matrix<-matrix(rep(3,N*M),nrow = N)
pca.hist<-sym.histogram.pca(Hardwood.histogram,BIN.Matrix)
Hardwood.quantiles.PCA<-quantiles.RSDA(pca.hist$sym.hist.matrix.PCA,3)

## End(Not run)
```

---

quantiles.RSDA.KS   *quantiles.RSDA.KS*

Description

quantiles.RSDA.KS

Usage

```
quantiles.RSDA.KS(histogram.matrix, num.quantiles)
```

Arguments

`histogram.matrix`
A matrix of histograms

`num.quantiles` Number of quantiles

Value

Quantiles of a Histogram Matrix

Author(s)

Jorge Arce Garro

Examples

```
## Not run:
data("hardwoodBrito")
Hardwood.histogram<-hardwoodBrito
Hardwood.cols<-colnames(Hardwood.histogram)
Hardwood.names<-row.names(Hardwood.histogram)
 M<-length(Hardwood.cols)
 N<-length(Hardwood.names)
 BIN.Matrix<-matrix(rep(3,N*M),nrow = N)
pca.hist<-sym.histogram.pca(Hardwood.histogram,BIN.Matrix)
quantiles.RSDA.KS<-quantiles.RSDA(pca.hist$sym.hist.matrix.PCA,100)

## End(Not run)
```

R2.L *Lower boundary correlation coefficient.*

Description

Compute the lower boundary correlation coefficient for two interval variables.

Usage

```
R2.L(ref, pred)
```

Arguments

ref	Variable that was predicted.
pred	The prediction given by the model.

Value

The lower boundary correlation coefficient.

Author(s)

Oldemar Rodriguez Rojas

References

LIMA-NETO, E.A., DE CARVALHO, F.A.T., (2008). Centre and range method to fitting a linear regression model on symbolic interval data. Computational Statistics and Data Analysis 52, 1500-1515.

LIMA-NETO, E.A., DE CARVALHO, F.A.T., (2010). Constrained linear regression models for symbolic interval-valued variables. Computational Statistics and Data Analysis 54, 333-347.

See Also

sym.glm

Examples

```
data(int_prost_train)
data(int_prost_test)
res.cm <- sym.lm(lpsa ~ ., sym.data = int_prost_train, method = "cm")
pred.cm <- sym.predict(res.cm, int_prost_test)
R2.L(int_prost_test$lpsa, pred.cm$Fitted)
```

R2.U *Upper boundary correlation coefficient.*

Description

Compute the upper boundary correlation coefficient for two interval variables.

Usage

```
R2.U(ref, pred)
```

Arguments

ref	Variable that was predicted.
pred	The prediction given by the model.

Value

The upper boundary correlation coefficient.

Author(s)

Oldemar Rodriguez Rojas

References

LIMA-NETO, E.A., DE CARVALHO, F.A.T., (2008). Centre and range method to fitting a linear regression model on symbolic interval data. Computational Statistics and Data Analysis 52, 1500-1515.

LIMA-NETO, E.A., DE CARVALHO, F.A.T., (2010). Constrained linear regression models for symbolic interval-valued variables. Computational Statistics and Data Analysis 54, 333-347.

See Also

sym.glm

Examples

```
data(int_prost_train)
data(int_prost_test)
res.cm <- sym.lm(lpsa ~ ., sym.data = int_prost_train, method = "cm")
pred.cm <- sym.predict(res.cm, int_prost_test)
R2.U(int_prost_test$lpsa, pred.cm$Fitted)
```

`read.sym.table` *Read a Symbolic Table*

Description

It reads a symbolic data table from a CSV file.

Usage

```
read.sym.table(file, header = TRUE, sep, dec, row.names = NULL)
```

Arguments

`file`	The name of the CSV file.
`header`	As in R function read.table
`sep`	As in R function read.table
`dec`	As in R function read.table
`row.names`	As in R function read.table

Details

The labels $C means that follows a continuous variable, $I means an interval variable, $H means a histogram variables and $S means set variable. In the first row each labels should be follow of a name to variable and to the case of histogram a set variables types the names of the modalities (categories) . In data rows for continuous variables we have just one value, for interval variables we have the minimum and the maximum of the interval, for histogram variables we have the number of modalities and then the probability of each modality and for set variables we have the cardinality of the set and next the elements of the set.

The format is the CSV file should be like:

$C F1 $I F2 F2 $H F3 M1 M2 M3 $S F4 E1 E2 E3 E4

Case1 $C 2.8 $I 1 2 $H 3 0.1 0.7 0.2 $S 4 e g k i

Case2 $C 1.4 $I 3 9 $H 3 0.6 0.3 0.1 $S 4 a b c d

Case3 $C 3.2 $I -1 4 $H 3 0.2 0.2 0.6 $S 4 2 1 b c

Case4 $C -2.1 $I 0 2 $H 3 0.9 0.0 0.1 $S 4 3 4 c a

Case5 $C -3.0 $I -4 -2 $H 3 0.6 0.0 0.4 $S 4 e i g k

The internal format is:
$N
[1] 5
$M
[1] 4
$sym.obj.names
[1] 'Case1' 'Case2' 'Case3' 'Case4' 'Case5'
$sym.var.names
[1] 'F1' 'F2' 'F3' 'F4'
$sym.var.types
[1] '$C' '$I' '$H' '$S'
$sym.var.length
[1] 1 2 3 4
$sym.var.starts
[1] 2 4 8 13
$meta
$C F1 $I F2 F2 $H F3 M1 M2 M3 $S F4 E1 E2 E3 E4
Case1 $C 2.8 $I 1 2 $H 3 0.1 0.7 0.2 $S 4 e g k i
Case2 $C 1.4 $I 3 9 $H 3 0.6 0.3 0.1 $S 4 a b c d
Case3 $C 3.2 $I -1 4 $H 3 0.2 0.2 0.6 $S 4 2 1 b c
Case4 $C -2.1 $I 0 2 $H 3 0.9 0.0 0.1 $S 4 3 4 c a
Case5 $C -3.0 $I -4 -2 $H 3 0.6 0.0 0.4 $S 4 e i g k
$data
F1 F2 F2.1 M1 M2 M3 E1 E2 E3 E4
Case1 2.8 1 2 0.1 0.7 0.2 e g k i
Case2 1.4 3 9 0.6 0.3 0.1 a b c d
Case3 3.2 -1 4 0.2 0.2 0.6 2 1 b c
Case4 -2.1 0 2 0.9 0.0 0.1 3 4 c a
Case5 -3.0 -4 -2 0.6 0.0 0.4 e i g k

Value

Return a symbolic data table structure.

Author(s)

Oldemar Rodriguez Rojas

References

Bock H-H. and Diday E. (eds.) (2000). Analysis of Symbolic Data. Exploratory methods for extracting statistical information from complex data. Springer, Germany.

See Also

display.sym.table

Examples

```
## Not run:
data(example1)
write.sym.table(example1,
  file = "temp4.csv", sep = "|", dec = ".", row.names = TRUE,
  col.names = TRUE
)
ex1 <- read.sym.table("temp4.csv", header = TRUE, sep = "|", dec = ".", row.names = 1)

## End(Not run)
```

RMSE.L *Lower boundary root-mean-square error*

Description

Compute the lower boundary root-mean-square error.

Usage

```
RMSE.L(ref, pred)
```

Arguments

ref	Variable that was predicted.
pred	The prediction given by the model.

Value

The lower boundary root-mean-square error.

Author(s)

Oldemar Rodriguez Rojas.

References

LIMA-NETO, E.A., DE CARVALHO, F.A.T., (2008). Centre and range method to fitting a linear regression model on symbolic interval data. Computational Statistics and Data Analysis 52, 1500-1515.

LIMA-NETO, E.A., DE CARVALHO, F.A.T., (2010). Constrained linear regression models for symbolic interval-valued variables. Computational Statistics and Data Analysis 54, 333-347.

See Also

sym.glm

`RMSE.U` *Upper boundary root-mean-square error*

Description

Compute the upper boundary root-mean-square error.

Usage

```
RMSE.U(ref, pred)
```

Arguments

`ref`	Variable that was predicted.
`pred`	The prediction given by the model.

Value

The upper boundary root-mean-square error.

Author(s)

Oldemar Rodriguez Rojas

References

LIMA-NETO, E.A., DE CARVALHO, F.A.T., (2008). Centre and range method to fitting a linear regression model on symbolic interval data. Computational Statistics and Data Analysis 52, 1500-1515.

LIMA-NETO, E.A., DE CARVALHO, F.A.T., (2010). Constrained linear regression models for symbolic interval-valued variables. Computational Statistics and Data Analysis 54, 333-347.

See Also

sym.glm

RSDA *R to Symbolic Data Analysis*

Description

This work is framed inside the Symbolic Data Analysis (SDA). The objective of this work is to implement in R to the symbolic case certain techniques of the automatic classification, as well as some lineal models. These implementations will always be made following two fundamental principles in Symbolic Data Analysis like they are: Classic Data Analysis should always be a case particular case of the Symbolic Data Analysis and both, the exit as the input in an Symbolic Data Analysis should be symbolic. We implement for variables of type interval the mean, the median, the mean of the extreme values, the standard deviation, the deviation quartil, the dispersion boxes and the correlation also three new methods are also presented to carry out the lineal regression for variables of type interval. We also implement in this R package the method of Principal Components Analysis in two senses: First, we propose three ways to project the interval variables in the circle of correlations in such way that is reflected the variation or the inexactness of the variables. Second, we propose an algorithm to make the Principal Components Analysis for variables of type histogram. We implement a method for multidimensional scaling of interval data, denominated INTERSCAL.

Details

Package:	RSDA
Type:	Package
Version:	3.1.0
Date:	2023-04-21
License:	GPL (>=2)

Most of the function of the package stars from a symbolic data table that can be store in a CSV file withe follwing forma: In the first row the labels $C means that follows a continuous variable, $I means an interval variable, $H means a histogram variables and $S means set variable. In the first row each labels should be follow of a name to variable and to the case of histogram a set variables types the names of the modalities (categories) . In data rows for continuous variables we have just one value, for interval variables we have the minimum and the maximum of the interval, for histogram variables we have the number of modalities and then the probability of each modality and for set variables we have the cardinality of the set and next the elements of the set.

Author(s)

Oldemar Rodriguez Rojas
Maintainer: Oldemar Rodriguez Rojas <oldemar.rodriguez@ucr.ac.cr>

References

Billard L. and Diday E. (2006). Symbolic data analysis: Conceptual statistics and data mining. Wiley, Chichester.

Billard L., Douzal-Chouakria A. and Diday E. (2011) Symbolic Principal Components For Interval-Valued Observations, Statistical Analysis and Data Mining. 4 (2), 229-246. Wiley.

Bock H-H. and Diday E. (eds.) (2000). Analysis of Symbolic Data. Exploratory methods for extracting statistical information from complex data. Springer, Germany.

Carvalho F., Souza R.,Chavent M., and Lechevallier Y. (2006) Adaptive Hausdorff distances and dynamic clustering of symbolic interval data. Pattern Recognition Letters Volume 27, Issue 3, February 2006, Pages 167-179

Cazes P., Chouakria A., Diday E. et Schektman Y. (1997). Extension de l'analyse en composantes principales a des donnees de type intervalle, Rev. Statistique Appliquee, Vol. XLV Num. 3 pag. 5-24, France.

Diday, E., Rodriguez O. and Winberg S. (2000). Generalization of the Principal Components Analysis to Histogram Data, 4th European Conference on Principles and Practice of Knowledge Discovery in Data Bases, September 12-16, 2000, Lyon, France.

Chouakria A. (1998) Extension des methodes d'analysis factorialle a des donnees de type intervalle, Ph.D. Thesis, Paris IX Dauphine University.

Makosso-Kallyth S. and Diday E. (2012). Adaptation of interval PCA to symbolic histogram variables, Advances in Data Analysis and Classification July, Volume 6, Issue 2, pp 147-159. Rodriguez, O. (2000). Classification et Modeles Lineaires en Analyse des Donnees Symboliques. Ph.D. Thesis, Paris IX-Dauphine University.

`sd` *Generic function for the standard desviation*

Description

Compute the symbolic standard desviation.

Usage

```
sd(x, ...)

## Default S3 method:
sd(x, na.rm = FALSE, ...)

## S3 method for class 'symbolic_interval'
sd(x, method = c("centers", "interval", "billard"), na.rm = FALSE, ...)

## S3 method for class 'symbolic_tbl'
sd(x, ...)
```

Arguments

`x`	A symbolic variable.
`...`	As in R sd function.
`na.rm`	As in R sd function.
`method`	The method to be use.

Value

return a real number.

Author(s)

Oldemar Rodriguez Rojas

References

Billard L. and Diday E. (2006). Symbolic data analysis: Conceptual statistics and data mining. Wiley, Chichester.

Rodriguez, O. (2000). Classification et Modeles Lineaires en Analyse des Donnees Symboliques. Ph.D. Thesis, Paris IX-Dauphine University.

SDS.to.RSDA *SDS SODAS files to RSDA files.*

Description

To convert SDS SODAS files to RSDA files.

Usage

```
SDS.to.RSDA(file.path, labels = FALSE)
```

Arguments

`file.path`	Disk path where the SODAS *.SDA file is.
`labels`	If we want to include SODAS SDA files lebels in RSDA file.

Value

A RSDA symbolic data file.

Author(s)

Olger Calderon and Roberto Zuniga.

References

Bock H-H. and Diday E. (eds.) (2000). Analysis of Symbolic Data. Exploratory methods for extracting statistical information from complex data. Springer, Germany.

See Also

SODAS.to.RSDA

Examples

```
## Not run:
# We can read the file directly from the SODAS SDA file as follows:
# We can save the file in CSV to RSDA format as follows:
setwd('C:/Program Files (x86)/DECISIA/SODAS version 2.0/bases/')
result <- SDS.to.RSDA(file.path='hani3101.sds')
# We can save the file in CSV to RSDA format as follows:
write.sym.table(result, file='hani3101.csv', sep=';',dec='.', row.names=TRUE,

## End(Not run)
```

SODAS.to.RSDA *XML SODAS files to RSDA files.*

Description

To convert XML SODAS files to RSDA files.

Usage

```
SODAS.to.RSDA(XMLPath, labels = T)
```

Arguments

XMLPath	Disk path where the SODAS *.XML file is.
labels	If we want to include SODAS XML files lebels in RSDA file.

Value

A RSDA symbolic data file.

Author(s)

Olger Calderon and Roberto Zuniga.

References

Bock H-H. and Diday E. (eds.) (2000). Analysis of Symbolic Data. Exploratory methods for extracting statistical information from complex data. Springer, Germany.

See Also

SDS.to.RSDA

Examples

```
## Not run:
# We can read the file directly from the SODAS XML file as follows:
# abalone<-SODAS.to.RSDA('C:/Program Files (x86)/DECISIA/SODAS version 2.0/bases/abalone.xml)
# We can save the file in CSV to RSDA format as follows:
# write.sym.table(sodas.ex1, file='abalone.csv', sep=';',dec='.', row.names=TRUE,
#                col.names=TRUE)
# We read the file from the CSV file,
# this is not necessary if the file is read directly from
# XML using SODAS.to.RSDA as in the first statement in this example.
data(abalone)
res <- sym.interval.pca(abalone, "centers")
sym.scatterplot(sym.var(res$Sym.Components, 1), sym.var(res$Sym.Components, 2),
  labels = TRUE, col = "red", main = "PCA Oils Data"
)
sym.scatterplot3d(sym.var(res$Sym.Components, 1), sym.var(res$Sym.Components, 2),
  sym.var(res$Sym.Components, 3),
  color = "blue", main = "PCA Oils Data"
)
sym.scatterplot.ggplot(sym.var(res$Sym.Components, 1), sym.var(res$Sym.Components, 2),
  labels = TRUE
)
sym.circle.plot(res$Sym.Prin.Correlations)

## End(Not run)
```

`stand.data` *Standardized Intervals*

Description

Standardized Intervals

Usage

```
stand.data(sym.data, data.mean, data.stan, nn, mm)
```

Arguments

`sym.data`	An Interval Matrix
`data.mean`	A vector of means
`data.stan`	A vector of standard deviation
`nn`	Number of concepts
`mm`	Number of variables

Value

Standardized intervals

`sym.all.quantiles.mesh3D.plot`
sym.all.quantiles.mesh3D.plot

Description

sym.all.quantiles.mesh3D.plot

Usage

```
sym.all.quantiles.mesh3D.plot(
  quantiles.sym,
  concept.names,
  var.names,
  Title,
  axes.x.label,
  axes.y.label,
  label.name
)
```

Arguments

`quantiles.sym`	A quantile matrix
`concept.names`	Concept Names
`var.names`	Variables to plot
`Title`	Plot title
`axes.x.label`	Label of axis X
`axes.y.label`	Label of axis Y
`label.name`	Concept Variable

Value

3D Mesh Plot

Author(s)

Jorge Arce Garro

Examples

```
## Not run:
 data("hardwoodBrito")
 Hardwood.histogram<-hardwoodBrito
 Hardwood.cols<-colnames(Hardwood.histogram)
 Hardwood.names<-row.names(Hardwood.histogram)
 M<-length(Hardwood.cols)
```

```
N<-length(Hardwood.names)
BIN.Matrix<-matrix(rep(3,N*M),nrow = N)
pca.hist<-sym.histogram.pca(Hardwood.histogram,BIN.Matrix)
Hardwood.quantiles.PCA<-quantiles.RSDA(pca.hist$sym.hist.matrix.PCA,3)
label.name<-"Hard Wood"
Title<-"First Principal Plane"
axes.x.label<- "First Principal Component (84.83%)"
axes.y.label<- "Second Principal Component (9.70%)"
concept.names<-c("ACER")
var.names<-c("PC.1","PC.2")
concept.names<-row.names(Hardwood.quantiles.PCA)
sym.all.quantiles.mesh3D.plot(Hardwood.quantiles.PCA,
                              concept.names,
                              var.names,
                              Title,
                              axes.x.label,
                              axes.y.label,
                              label.name)

## End(Not run)
```

`sym.all.quantiles.plot` *sym.all.quantiles.plot*

Description

sym.all.quantiles.plot

Usage

```
sym.all.quantiles.plot(
  quantiles.sym,
  concept.names,
  var.names,
  Title,
  axes.x.label,
  axes.y.label,
  label.name
)
```

Arguments

`quantiles.sym`	A quantile matrix
`concept.names`	Concept Names
`var.names`	Variables to plot
`Title`	Plot title
`axes.x.label`	Label of axis X

axes.y.label Label of axis Y

label.name Concept Variable

Value

3D Scatter Plot

Author(s)

Jorge Arce Garro

Examples

```
## Not run:
 data("hardwoodBrito")
 Hardwood.histogram<-hardwoodBrito
 Hardwood.cols<-colnames(Hardwood.histogram)
 Hardwood.names<-row.names(Hardwood.histogram)
 M<-length(Hardwood.cols)
 N<-length(Hardwood.names)
 BIN.Matrix<-matrix(rep(3,N*M),nrow = N)
 pca.hist<-sym.histogram.pca(Hardwood.histogram,BIN.Matrix)
 Hardwood.quantiles.PCA<-quantiles.RSDA(pca.hist$sym.hist.matrix.PCA,3)
 label.name<-"Hard Wood"
 Title<-"First Principal Plane"
 axes.x.label<- "First Principal Component (84.83%)"
 axes.y.label<- "Second Principal Component (9.70%)"
 concept.names<-c("ACER")
 var.names<-c("PC.1","PC.2")

 concept.names<-row.names(Hardwood.quantiles.PCA)
 sym.all.quantiles.plot(Hardwood.quantiles.PCA,
                           concept.names,
                           var.names,
                           Title,
                           axes.x.label,
                           axes.y.label,
                           label.name)

## End(Not run)
```

---

sym.circle.plot *Symbolic Circle of Correlations*

Description

Plot the symbolic circle of correlations.

Usage

```
sym.circle.plot(prin.corre)
```

Arguments

prin.corre	A symbolic interval data matrix with correlations between the variables and the principals componets, both of interval type.

Value

Plot the symbolic circle

Author(s)

Oldemar Rodriguez Rojas

References

Rodriguez O. (2012). The Duality Problem in Interval Principal Components Analysis. The 3rd Workshop in Symbolic Data Analysis, Madrid.

Examples

```
data(oils)
res <- sym.pca(oils, "centers")
sym.circle.plot(res$Sym.Prin.Correlations)
```

---

sym.dist.interval *Distance for Symbolic Interval Variables.*

Description

This function computes and returns the distance matrix by using the specified distance measure to compute distance between symbolic interval variables.

Usage

```
sym.dist.interval(
  sym.data,
  gamma = 0.5,
  method = "Minkowski",
  normalize = TRUE,
  SpanNormalize = FALSE,
  q = 1,
  euclidea = TRUE,
  pond = rep(1, length(variables))
)
```

Arguments

sym.data	A symbolic object
gamma	gamma value for the methods ichino and minkowski.
method	Method to use (Gowda.Diday, Ichino, Minkowski, Hausdorff)
normalize	A logical value indicating whether normalize the data in the ichino or hausdorff method.
SpanNormalize	A logical value indicating whether
q	q value for the hausdorff method.
euclidea	A logical value indicating whether use the euclidean distance.
pond	A numeric vector
variables	Numeric vector with the number of the variables to use.

Value

An object of class 'dist'

---

sym.gbm *Generalized Boosted Symbolic Regression*

Description

Generalized Boosted Symbolic Regression

Usage

```
sym.gbm(
  formula,
  sym.data,
  method = c("cm", "crm"),
  distribution = "gaussian",
  interaction.depth = 1,
  n.trees = 500,
  shrinkage = 0.1
)
```

Arguments

formula	A symbolic description of the model to be fit. The formula may include an offset term (e.g. y~offset(n)+x). If keep.data = FALSE in the initial call to gbm then it is the user's responsibility to resupply the offset to gbm.more.
sym.data	symbolic data table
method	cm crm
distribution	distribution

interaction.depth
: Integer specifying the maximum depth of each tree (i.e., the highest level of variable interactions allowed). A value of 1 implies an additive model, a value of 2 implies a model with up to 2-way interactions, etc. Default is 1.

n.trees
: Integer specifying the total number of trees to fit. This is equivalent to the number of iterations and the number of basis functions in the additive expansion. Default is 100.

shrinkage
: A shrinkage parameter applied to each tree in the expansion. Also known as the learning rate or step-size reduction; 0.001 to 0.1 usually work, but a smaller learning rate typically requires more trees. Default is 0.1.

References

Lima-Neto, E.A., De Carvalho, F.A.T., (2008). Centre and range method to fitting a linear regression model on symbolic interval data. Computational Statistics and Data Analysis52, 1500-1515

Lima-Neto, E.A., De Carvalho, F.A.T., (2010). Constrained linear regression models for symbolic interval-valued variables. Computational Statistics and Data Analysis 54, 333-347

Lima Neto, E.d.A., de Carvalho, F.d.A.T. Nonlinear regression applied to interval-valued data. Pattern Anal Applic 20, 809–824 (2017). https://doi.org/10.1007/s10044-016-0538-y

Rodriguez, O. (2018). Shrinkage linear regression for symbolic interval-valued variables.Journal MODULAD 2018, vol. Modulad 45, pp.19-38

---

sym.glm *Lasso, Ridge and and Elastic Net Linear regression model to interval variables*

Description

Execute Lasso, Ridge and and Elastic Net Linear regression model to interval variables.

Usage

```
sym.glm(sym.data, response = 1, method = c('cm', 'crm'),
alpha = 1, nfolds = 10, grouped = TRUE)
```

Arguments

sym.data
: Should be a symbolic data table read with the function read.sym.table(...).

response
: The number of the column where is the response variable in the interval data table.

method
: 'cm' to generalized Center Method and 'crm' to generalized Center and Range Method.

alpha
: alpha=1 is the lasso penalty, and alpha=0 the ridge penalty. 0<alpha<1 is the elastic net method.

nfolds	Number of folds - default is 10. Although nfolds can be as large as the sample size (leave-one-out CV), it is not recommended for large datasets. Smallest value allowable is nfolds=3
grouped	This is an experimental argument, with default TRUE, and can be ignored by most users.

Value

An object of class 'cv.glmnet' is returned, which is a list with the ingredients of the cross-validation fit.

Author(s)

Oldemar Rodriguez Rojas

References

Rodriguez O. (2013). A generalization of Centre and Range method for fitting a linear regression model to symbolic interval data using Ridge Regression, Lasso and Elastic Net methods. The IFCS2013 conference of the International Federation of Classification Societies, Tilburg University Holland.

See Also

sym.lm

---

sym.histogram *Create an symbolic_histogram type object*

Description

Create an symbolic_histogram type object

Usage

```
sym.histogram(x = double(), breaks = NA_real_)
```

Arguments

x	character vector
breaks	a vector giving the breakpoints between histogram cells

Value

a symbolic histogram

Examples

```
sym.histogram(iris$Sepal.Length)
```

sym.histogram.pca *sym.histogram.pca*

Description

sym.histogram.pca

Usage

```
sym.histogram.pca(sym.hist.matrix, BIN.Matrix, method = NULL)
```

Arguments

`sym.hist.matrix`	A Histogram matrix
`BIN.Matrix`	A matrix with the number of bins for each individual and variable
`method`	Weigthed Method

Value

Histogram PCA

Author(s)

Jorge Arce Garro

Examples

```
## Not run:
data("hardwoodBrito")
Hardwood.histogram<-hardwoodBrito
weighted.center<-weighted.center.Hist.RSDA(Hardwood.histogram)
 M<-length(Hardwood.cols)
 N<-length(Hardwood.names)
 BIN.Matrix<-matrix(rep(3,N*M),nrow = N)
pca.hist<-sym.histogram.pca(Hardwood.histogram,BIN.Matrix)
pca.hist

## End(Not run)
```

sym.interval *Create an symbolic_interval type object*

Description

Create an symbolic_interval type object

Usage

```
sym.interval(x = numeric(), .min = min, .max = max)
```

Arguments

x	numeric vector
.min	function that will be used to calculate the minimum interval
.max	function that will be used to calculate the maximum interval

Value

a symbolic interval

Examples

```
sym.interval(c(1, 2, 4, 5))
sym.interval(1:10)
```

sym.interval.pc *Compute a symbolic interval principal components curves*

Description

Compute a symbolic interval principal components curves

Usage

```
sym.interval.pc(sym.data, method = c('vertex', 'centers'), maxit, plot, scale, center)
```

Arguments

sym.data	Shoud be a symbolic data table read with the function read.sym.table(...)
method	It should be 'vertex' or 'centers'.
maxit	Maximum number of iterations.
plot	TRUE to plot immediately, FALSE if you do not want to plot.
scale	TRUE to standardize the data.
center	TRUE to center the data.

Value

prin.curve: This a symbolic data table with the interval principal components. As this is a symbolic data table we can apply over this table any other symbolic data analysis method (symbolic propagation).

cor.ps: This is the interval correlations between the original interval variables and the interval principal components, it can be use to plot the symbolic circle of correlations.

Author(s)

Jorge Arce.

References

Arce J. and Rodriguez O. (2015) 'Principal Curves and Surfaces to Interval Valued Variables'. The 5th Workshop on Symbolic Data Analysis, SDA2015, Orleans, France, November.

Hastie,T. (1984). Principal Curves and Surface. Ph.D Thesis Stanford University.

Hastie,T. & Weingessel,A. (2014). princurve - Fits a Principal Curve in Arbitrary Dimension.R package version 1.1–12 http://cran.r-project.org/web/packages/princurve/index.html.

Hastie,T. & Stuetzle, W. (1989). Principal Curves. Journal of the American Statistical Association, Vol. 84-406, 502–516.

Hastie, T., Tibshirani, R. & Friedman, J. (2008). The Elements of Statistical Learning; Data Mining, Inference and Prediction. Springer, New York.

See Also

sym.interval.pca

Examples

```
## Not run:
data(oils)
res.vertex.ps <- sym.interval.pc(oils, "vertex", 150, FALSE, FALSE, TRUE)
class(res.vertex.ps$sym.prin.curve) <- c("sym.data.table")
sym.scatterplot(res.vertex.ps$sym.prin.curve[, 1], res.vertex.ps$sym.prin.curve[, 2],
  labels = TRUE, col = "red", main = "PSC Oils Data"
)

data(facedata)
res.vertex.ps <- sym.interval.pc(facedata, "vertex", 150, FALSE, FALSE, TRUE)
class(res.vertex.ps$sym.prin.curve) <- c("sym.data.table")
sym.scatterplot(res.vertex.ps$sym.prin.curve[, 1], res.vertex.ps$sym.prin.curve[, 2],
  labels = TRUE, col = "red", main = "PSC Face Data"
)

## End(Not run)
```

sym.interval.pc.limits *Symbolic interval principal curves limits*

Description

Symbolic interval principal curves limits.

Usage

```
sym.interval.pc.limits(sym.data, prin.curve, num.vertex, lambda, var.ord)
```

Arguments

`sym.data`	Symbolic interval data table.
`prin.curve`	Principal curves.
`num.vertex`	Number of vertices of the hipercube.
`lambda`	Lambda.
`var.ord`	Order of the variables.

Author(s)

Jorge Arce.

References

Arce J. and Rodriguez O. (2015) 'Principal Curves and Surfaces to Interval Valued Variables'. The 5th Workshop on Symbolic Data Analysis, SDA2015, Orleans, France, November.

Hastie,T. (1984). Principal Curves and Surface. Ph.D Thesis Stanford University.

Hastie,T. & Weingessel,A. (2014). princurve - Fits a Principal Curve in Arbitrary Dimension.R package version 1.1–12 http://cran.r-project.org/web/packages/princurve/index.html.

Hastie,T. & Stuetzle, W. (1989). Principal Curves. Journal of the American Statistical Association, Vol. 84-406, 502–516.

Hastie, T., Tibshirani, R. & Friedman, J. (2008). The Elements of Statistical Learning; Data Mining, Inference and Prediction. Springer, New York.

See Also

sym.interval.pc

`sym.kmeans` *Symbolic k-Means*

Description

This is a function is to carry out a k-means overs a interval symbolic data matrix.

Usage

```
sym.kmeans(sym.data, k = 3, iter.max = 10, nstart = 1,
algorithm = c('Hartigan-Wong', 'Lloyd', 'Forgy', 'MacQueen'))
```

Arguments

`sym.data`	Symbolic data table.
`k`	The number of clusters.
`iter.max`	Maximun number of iterations.
`nstart`	As in R kmeans function.
`algorithm`	The method to be use, as in kmeans R function.

Value

This function return the following information:

K-means clustering with 3 clusters of sizes 2, 2, 4

Cluster means:

GRA FRE IOD SAP

1 0.93300 -13.500 193.500 174.75

2 0.86300 30.500 54.500 195.25

3 0.91825 -6.375 95.375 191.50

Clustering vector:

L P Co S Ca O B H

1 1 3 3 3 3 2 2

Within cluster sum of squares by cluster:

[1] 876.625 246.125 941.875

(between_SS / total_SS = 92.0

Available components:

[1] 'cluster' 'centers' 'totss' 'withinss' 'tot.withinss' 'betweenss'

[7] 'size'

Author(s)

Oldemar Rodriguez Rojas

References

Carvalho F., Souza R.,Chavent M., and Lechevallier Y. (2006) Adaptive Hausdorff distances and dynamic clustering of symbolic interval data. Pattern Recognition Letters Volume 27, Issue 3, February 2006, Pages 167-179

See Also

sym.hclust

Examples

```
data(oils)
sk <- sym.kmeans(oils, k = 3)
sk$cluster
```

`sym.knn` *Symbolic k-Nearest Neighbor Regression*

Description

Symbolic k-Nearest Neighbor Regression

Usage

```
sym.knn(
  formula,
  sym.data,
  method = c("cm", "crm"),
  scale = TRUE,
```

```
  kmax = 20,
  kernel = "triangular"
)
```

Arguments

`formula`	a formula object.
`sym.data`	symbolc data.table
`method`	cm or crm
`scale`	logical, scale variable to have equal sd.
`kmax`	maximum number of k, if ks is not specified.
`kernel`	kernel to use. Possible choices are "rectangular" (which is standard unweighted knn), "triangular", "epanechnikov" (or beta(2,2)), "biweight" (or beta(3,3)), "triweight" (or beta(4,4)), "cos", "inv", "gaussian" and "optimal".

References

Lima-Neto, E.A., De Carvalho, F.A.T., (2008). Centre and range method to fitting a linear regression model on symbolic interval data. Computational Statistics and Data Analysis52, 1500-1515

Lima-Neto, E.A., De Carvalho, F.A.T., (2010). Constrained linear regression models for symbolic interval-valued variables. Computational Statistics and Data Analysis 54, 333-347

Lima Neto, E.d.A., de Carvalho, F.d.A.T. Nonlinear regression applied to interval-valued data. Pattern Anal Applic 20, 809–824 (2017). https://doi.org/10.1007/s10044-016-0538-y

Rodriguez, O. (2018). Shrinkage linear regression for symbolic interval-valued variables.Journal MODULAD 2018, vol. Modulad 45, pp.19-38

`sym.lm` *CM and CRM Linear regression model.*

Description

To execute the Center Method (CR) and Center and Range Method (CRM) to Linear regression.

Usage

```
sym.lm(formula, sym.data, method = c('cm', 'crm'))
```

Arguments

`formula`	An object of class 'formula' (or one that can be coerced to that class): a symbolic description of the model to be fitted.
`sym.data`	Should be a symbolic data table read with the function read.sym.table(...).
`method`	'cm' to Center Method and 'crm' to Center and Range Method.

Details

Models for lm are specified symbolically. A typical model has the form response ~ terms where response is the (numeric) response vector and terms is a series of terms which specifies a linear predictor for response. A terms specification of the form first + second indicates all the terms in first together with all the terms in second with duplicates removed. A specification of the form first:second indicates the set of terms obtained by taking the interactions of all terms in first with all terms in second. The specification first*second indicates the cross of first and second. This is the same as first + second + first:second.

Value

sym.lm returns an object of class 'lm' or for multiple responses of class c('mlm', 'lm')

Author(s)

Oldemar Rodriguez Rojas

References

LIMA-NETO, E.A., DE CARVALHO, F.A.T., (2008). Centre and range method to fitting a linear regression model on symbolic interval data. Computational Statistics and Data Analysis 52, 1500-1515.

LIMA-NETO, E.A., DE CARVALHO, F.A.T., (2010). Constrained linear regression models for symbolic interval-valued variables. Computational Statistics and Data Analysis 54, 333-347.

Examples

```
data(int_prost_train)
data(int_prost_test)
res.cm <- sym.lm(lpsa ~ ., sym.data = int_prost_train, method = "cm")
res.cm
```

`sym.mcfa` *sym.mcfa*

Description

This function executes a Multiple Correspondence Factor Analysis for variables of set type.

Usage

```
sym.mcfa(sym.data, pos.var)
```

Arguments

`sym.data`	A symbolic data table containing at least two set type variables.
`pos.var`	Column numbers in the symbolic data table that contain the set type variables.

Author(s)

Jorge Arce

References

Arce J. and Rodriguez, O. (2018). Multiple Correspondence Analysis for Symbolic Multi–Valued Variables. On the Symbolic Data Analysis Workshop SDA 2018.

Benzecri, J.P. (1973). L' Analyse des Données. Tomo 2: L'Analyse des Correspondances. Dunod, Paris.

Castillo, W. and Rodriguez O. (1997). Algoritmo e implementacion del analisis factorial de correspondencias. Revista de Matematicas: Teoria y Aplicaciones, 24-31.

Takagi I. and Yadosiha H. (2011). Correspondence Analysis for symbolic contingency tables base on interval algebra. Elsevier Procedia Computer Science, 6, 352-357.

Rodriguez, O. (2007). Correspondence Analysis for Symbolic Multi–Valued Variables. CARME 2007 (Rotterdam, The Netherlands), http://www.carme-n.org/carme2007.

Examples

```
data("ex_mcfa1")
sym.table <- classic.to.sym(ex_mcfa1,
  concept = suspect,
  hair = sym.set(hair),
  eyes = sym.set(eyes),
  region = sym.set(region)
)
sym.table
```

---

sym.modal *Create an symbolic_modal type object*

Description

Create an symbolic_modal type object

Usage

```
sym.modal(x = character())
```

Arguments

x	character vector

Value

a symbolic modal

Examples

```
sym.modal(factor(c("a", "b", "b", "l")))
```

---

sym.nnet *Symbolic neural networks regression*

Description

Symbolic neural networks regression

Usage

```
sym.nnet(
  formula,
  sym.data,
  method = c("cm", "crm"),
  hidden = c(10),
  threshold = 0.05,
  stepmax = 1e+05
)
```

Arguments

formula	a symbolic description of the model to be fitted.
sym.data	symbolic data.table
method	cm crm
hidden	a vector of integers specifying the number of hidden neurons (vertices) in each layer.
threshold	a numeric value specifying the threshold for the partial derivatives of the error function as stopping criteria.
stepmax	the maximum steps for the training of the neural network. Reaching this maximum leads to a stop of the neural network's training process.

References

Lima-Neto, E.A., De Carvalho, F.A.T., (2008). Centre and range method to fitting a linear regression model on symbolic interval data. Computational Statistics and Data Analysis52, 1500-1515

Lima-Neto, E.A., De Carvalho, F.A.T., (2010). Constrained linear regression models for symbolic interval-valued variables. Computational Statistics and Data Analysis 54, 333-347

Lima Neto, E.d.A., de Carvalho, F.d.A.T. Nonlinear regression applied to interval-valued data. Pattern Anal Applic 20, 809–824 (2017). https://doi.org/10.1007/s10044-016-0538-y

Rodriguez, O. (2018). Shrinkage linear regression for symbolic interval-valued variables.Journal MODULAD 2018, vol. Modulad 45, pp.19-38

sym.pca *Interval Principal Components Analysis.*

Description

Cazes, Chouakria, Diday and Schektman (1997) proposed the Centers and the Tops Methods to extend the well known principal components analysis method to a particular kind of symbolic objects characterized by multi–values variables of interval type.

Usage

```
sym.pca(sym.data, ...)

## S3 method for class 'symbolic_tbl'
sym.pca(
  sym.data,
 method = c("classic", "tops", "centers", "principal.curves", "optimized.distance",
    "optimized.variance", "fixed"),
  fixed.matrix = NULL,
  ...
)
```

Arguments

sym.data	Shoud be a symbolic data table
...	further arguments passed to or from other methods.
method	It is use so select the method, 'classic' execute a classical principal component analysis over the centers of the intervals, 'tops' to use the vertices algorithm and 'centers' to use the centers algorithm.
fixed.matrix	Classic Matrix. It is use when the method chosen is "fixed".

Value

Sym.Components: This a symbolic data table with the interval principal components. As this is a symbolic data table we can apply over this table any other symbolic data analysis method (symbolic propagation).

Sym.Prin.Correlations: This is the interval correlations between the original interval variables and the interval principal components, it can be use to plot the symbolic circle of correlations.

Author(s)

Oldemar Rodriguez Rojas

References

Bock H-H. and Diday E. (eds.) (2000). Analysis of Symbolic Data. Exploratory methods for extracting statistical information from complex data. Springer, Germany.

Cazes P., Chouakria A., Diday E. et Schektman Y. (1997). Extension de l'analyse en composantes principales a des donnees de type intervalle, Rev. Statistique Appliquee, Vol. XLV Num. 3 pag. 5-24, France.

Chouakria A. (1998) Extension des methodes d'analysis factorialle a des donnees de type intervalle, Ph.D. Thesis, Paris IX Dauphine University.

Makosso-Kallyth S. and Diday E. (2012). Adaptation of interval PCA to symbolic histogram variables, Advances in Data Analysis and Classification July, Volume 6, Issue 2, pp 147-159.

Rodriguez, O. (2000). Classification et Modeles Lineaires en Analyse des Donnees Symboliques. Ph.D. Thesis, Paris IX-Dauphine University.

See Also

sym.histogram.pca

Examples

```
## Not run:
data(oils)
res <- sym.pca(oils, "centers")

sym.scatterplot(res$Sym.Components[, 1], res$Sym.Components[, 1],
  labels = TRUE, col = "red", main = "PCA Oils Data"
)
sym.scatterplot3d(res$Sym.Components[, 1], res$Sym.Components[, 2],
  res$Sym.Components[, 3],
  color = "blue", main = "PCA Oils Data"
)
sym.scatterplot.ggplot(res$Sym.Components[, 1], res$Sym.Components[, 2],
  labels = TRUE
)
sym.circle.plot(res$Sym.Prin.Correlations)

res <- sym.pca(oils, "classic")
plot(res, choix = "ind")
plot(res, choix = "var")

data(lynne2)
res <- sym.pca(lynne2, "centers")

sym.scatterplot(res$Sym.Components[, 1], res$Sym.Components[, 2],
  labels = TRUE, col = "red", main = "PCA Lynne Data"
)
sym.scatterplot3d(res$Sym.Components[, 1], res$Sym.Components[, 2],
  res$Sym.Components[, 3],
  color = "blue", main = "PCA Lynne Data"
)
sym.scatterplot.ggplot(res$Sym.Components[, 1], res$Sym.Components[, 2],
```

```
  labels = TRUE
)
sym.circle.plot(res$Sym.Prin.Correlations)

data(StudentsGrades)
st <- StudentsGrades
s.pca <- sym.pca(st)
plot(s.pca, choix = "ind")
plot(s.pca, choix = "var")

## End(Not run)
```

Sym.PCA.Hist.PCA.k.plot
Sym.PCA.Hist.PCA.k.plot

Description

Sym.PCA.Hist.PCA.k.plot

Usage

```
Sym.PCA.Hist.PCA.k.plot(
  data.sym.df,
  title.graph,
  concepts.name,
  title.x,
  title.y,
  pca.axes
)
```

Arguments

data.sym.df	Bins's projections
title.graph	Plot title
concepts.name	Concepts names
title.x	Label of axis X
title.y	Label of axis Y
pca.axes	Principal Component

Value

Concepts projected onto the Principal component chosen

Author(s)

Jorge Arce Garro

Examples

```
## Not run:
data("hardwoodBrito")
Hardwood.histogram<-hardwoodBrito
Hardwood.cols<-colnames(Hardwood.histogram)
Hardwood.names<-row.names(Hardwood.histogram)
 M<-length(Hardwood.cols)
 N<-length(Hardwood.names)
 BIN.Matrix<-matrix(rep(3,N*M),nrow = N)
pca.hist<-sym.histogram.pca(Hardwood.histogram,BIN.Matrix)
Hardwood.quantiles.PCA<-quantiles.RSDA(pca.hist$sym.hist.matrix.PCA,3)
ACER.p1<-Sym.PCA.Hist.PCA.k.plot(data.sym.df = pca.hist$Bins.df,
                                    title.graph = " ",
                                    concepts.name = c("ACER"),
                                    title.x = "First Principal Component (84.83%)",
                                    title.y = "Frequency",
                                    pca.axes = 1)

ACER.p1

## End(Not run)
```

`sym.predict` *Predict method to CM and CRM regression model*

Description

To execute predict method the Center Method (CR) and Center and Range Method (CRM) to Linear regression.

Usage

```
sym.predict(model, ...)

## S3 method for class 'symbolic_lm_cm'
sym.predict(model, new.sym.data, ...)

## S3 method for class 'symbolic_lm_crm'
sym.predict(model, new.sym.data, ...)

## S3 method for class 'symbolic_glm_cm'
sym.predict(model, new.sym.data, response, ...)

## S3 method for class 'symbolic_glm_crm'
sym.predict(model, new.sym.data, response, ...)
```

Arguments

`model`	The output of lm method.
`...`	additional arguments affecting the predictions produced.
`new.sym.data`	Should be a symbolic data table read with the function read.sym.table(...).
`response`	The number of the column where is the response variable in the interval data table.

Value

sym.predict produces a vector of predictions or a matrix of predictions and bounds with column names fit, lwr, and upr if interval is set. For type = 'terms' this is a matrix with a column per term and may have an attribute 'constant'

Author(s)

Oldemar Rodriguez Rojas

References

LIMA-NETO, E.A., DE CARVALHO, F.A.T., (2008). Centre and range method to fitting a linear regression model on symbolic interval data. Computational Statistics and Data Analysis 52, 1500-1515.

LIMA-NETO, E.A., DE CARVALHO, F.A.T., (2010). Constrained linear regression models for symbolic interval-valued variables. Computational Statistics and Data Analysis 54, 333-347.

See Also

sym.glm

Examples

```
data(int_prost_train)
data(int_prost_test)
model <- sym.lm(lpsa ~ ., sym.data = int_prost_train, method = "cm")
pred.cm <- sym.predict(model, int_prost_test)
pred.cm
```

`sym.predict.symbolic_gbm_cm` *Predict model_gbm_cm model*

Description

Predict model_gbm_cm model

Usage

```
## S3 method for class 'symbolic_gbm_cm'
sym.predict(model, new.sym.data, n.trees = 500, ...)
```

Arguments

`model`	model
`new.sym.data`	new data
`n.trees`	Integer specifying the total number of trees to fit. This is equivalent to the number of iterations and the number of basis functions in the additive expansion. Default is 100.
`...`	optional parameters

References

Lima-Neto, E.A., De Carvalho, F.A.T., (2008). Centre and range method to fitting a linear regression model on symbolic interval data. Computational Statistics and Data Analysis52, 1500-1515

Lima-Neto, E.A., De Carvalho, F.A.T., (2010). Constrained linear regression models for symbolic interval-valued variables. Computational Statistics and Data Analysis 54, 333-347

Lima Neto, E.d.A., de Carvalho, F.d.A.T. Nonlinear regression applied to interval-valued data. Pattern Anal Applic 20, 809–824 (2017). https://doi.org/10.1007/s10044-016-0538-y

Rodriguez, O. (2018). Shrinkage linear regression for symbolic interval-valued variables.Journal MODULAD 2018, vol. Modulad 45, pp.19-38

---

sym.predict.symbolic_gbm_crm

Predict model_gbm_crm model

Description

Predict model_gbm_crm model

Usage

```
## S3 method for class 'symbolic_gbm_crm'
sym.predict(model, new.sym.data, n.trees = 500, ...)
```

Arguments

`model`	model
`new.sym.data`	new data
`n.trees`	Integer specifying the total number of trees to fit. This is equivalent to the number of iterations and the number of basis functions in the additive expansion. Default is 100.
`...`	optional parameters

References

Lima-Neto, E.A., De Carvalho, F.A.T., (2008). Centre and range method to fitting a linear regression model on symbolic interval data. Computational Statistics and Data Analysis52, 1500-1515

Lima-Neto, E.A., De Carvalho, F.A.T., (2010). Constrained linear regression models for symbolic interval-valued variables. Computational Statistics and Data Analysis 54, 333-347

Lima Neto, E.d.A., de Carvalho, F.d.A.T. Nonlinear regression applied to interval-valued data. Pattern Anal Applic 20, 809–824 (2017). https://doi.org/10.1007/s10044-016-0538-y

Rodriguez, O. (2018). Shrinkage linear regression for symbolic interval-valued variables.Journal MODULAD 2018, vol. Modulad 45, pp.19-38

`sym.predict.symbolic_knn_cm` *Predict model_knn_cm model*

Description

Predict model_knn_cm model

Usage

```
## S3 method for class 'symbolic_knn_cm'
sym.predict(model, new.sym.data, ...)
```

Arguments

`model`	model
`new.sym.data`	new data
`...`	optional parameters

References

Lima-Neto, E.A., De Carvalho, F.A.T., (2008). Centre and range method to fitting a linear regression model on symbolic interval data. Computational Statistics and Data Analysis52, 1500-1515

Lima-Neto, E.A., De Carvalho, F.A.T., (2010). Constrained linear regression models for symbolic interval-valued variables. Computational Statistics and Data Analysis 54, 333-347

Lima Neto, E.d.A., de Carvalho, F.d.A.T. Nonlinear regression applied to interval-valued data. Pattern Anal Applic 20, 809–824 (2017). https://doi.org/10.1007/s10044-016-0538-y

Rodriguez, O. (2018). Shrinkage linear regression for symbolic interval-valued variables.Journal MODULAD 2018, vol. Modulad 45, pp.19-38

sym.predict.symbolic_knn_crm

Predict model_knn_crm model

Description

Predict model_knn_crm model

Usage

```
## S3 method for class 'symbolic_knn_crm'
sym.predict(model, new.sym.data, ...)
```

Arguments

`model`	model
`new.sym.data`	new data
`...`	optional parameters

References

Lima-Neto, E.A., De Carvalho, F.A.T., (2008). Centre and range method to fitting a linear regression model on symbolic interval data. Computational Statistics and Data Analysis52, 1500-1515

Lima-Neto, E.A., De Carvalho, F.A.T., (2010). Constrained linear regression models for symbolic interval-valued variables. Computational Statistics and Data Analysis 54, 333-347

Lima Neto, E.d.A., de Carvalho, F.d.A.T. Nonlinear regression applied to interval-valued data. Pattern Anal Applic 20, 809–824 (2017). https://doi.org/10.1007/s10044-016-0538-y

Rodriguez, O. (2018). Shrinkage linear regression for symbolic interval-valued variables.Journal MODULAD 2018, vol. Modulad 45, pp.19-38

sym.predict.symbolic_nnet_cm

Predict nnet_cm model

Description

Predict nnet_cm model

Usage

```
## S3 method for class 'symbolic_nnet_cm'
sym.predict(model, new.sym.data, ...)
```

Arguments

`model`	model
`new.sym.data`	new data
`...`	optional parameters

References

Lima-Neto, E.A., De Carvalho, F.A.T., (2008). Centre and range method to fitting a linear regression model on symbolic interval data. Computational Statistics and Data Analysis52, 1500-1515

Lima-Neto, E.A., De Carvalho, F.A.T., (2010). Constrained linear regression models for symbolic interval-valued variables. Computational Statistics and Data Analysis 54, 333-347

Lima Neto, E.d.A., de Carvalho, F.d.A.T. Nonlinear regression applied to interval-valued data. Pattern Anal Applic 20, 809–824 (2017). https://doi.org/10.1007/s10044-016-0538-y

Rodriguez, O. (2018). Shrinkage linear regression for symbolic interval-valued variables.Journal MODULAD 2018, vol. Modulad 45, pp.19-38

`sym.predict.symbolic_nnet_crm` *Predict nnet_crm model*

Description

Predict nnet_crm model

Usage

```
## S3 method for class 'symbolic_nnet_crm'
sym.predict(model, new.sym.data, ...)
```

Arguments

`model`	model
`new.sym.data`	new data
`...`	optional parameters

References

Lima-Neto, E.A., De Carvalho, F.A.T., (2008). Centre and range method to fitting a linear regression model on symbolic interval data. Computational Statistics and Data Analysis52, 1500-1515

Lima-Neto, E.A., De Carvalho, F.A.T., (2010). Constrained linear regression models for symbolic interval-valued variables. Computational Statistics and Data Analysis 54, 333-347

Lima Neto, E.d.A., de Carvalho, F.d.A.T. Nonlinear regression applied to interval-valued data. Pattern Anal Applic 20, 809–824 (2017). https://doi.org/10.1007/s10044-016-0538-y

Rodriguez, O. (2018). Shrinkage linear regression for symbolic interval-valued variables.Journal MODULAD 2018, vol. Modulad 45, pp.19-38

---

sym.predict.symbolic_rf_cm
Predict rf_cm model

Description

Predict rf_cm model

Usage

```
## S3 method for class 'symbolic_rf_cm'
sym.predict(model, new.sym.data, ...)
```

Arguments

`model`	model
`new.sym.data`	new data
`...`	optional parameters

References

Lima-Neto, E.A., De Carvalho, F.A.T., (2008). Centre and range method to fitting a linear regression model on symbolic interval data. Computational Statistics and Data Analysis52, 1500-1515

Lima-Neto, E.A., De Carvalho, F.A.T., (2010). Constrained linear regression models for symbolic interval-valued variables. Computational Statistics and Data Analysis 54, 333-347

Lima Neto, E.d.A., de Carvalho, F.d.A.T. Nonlinear regression applied to interval-valued data. Pattern Anal Applic 20, 809–824 (2017). https://doi.org/10.1007/s10044-016-0538-y

Rodriguez, O. (2018). Shrinkage linear regression for symbolic interval-valued variables.Journal MODULAD 2018, vol. Modulad 45, pp.19-38

---

sym.predict.symbolic_rf_crm
Predict rf_crm model

Description

Predict rf_crm model

Usage

```
## S3 method for class 'symbolic_rf_crm'
sym.predict(model, new.sym.data, ...)
```

Arguments

`model`	model
`new.sym.data`	new data
`...`	optional parameters

References

Lima-Neto, E.A., De Carvalho, F.A.T., (2008). Centre and range method to fitting a linear regression model on symbolic interval data. Computational Statistics and Data Analysis52, 1500-1515

Lima-Neto, E.A., De Carvalho, F.A.T., (2010). Constrained linear regression models for symbolic interval-valued variables. Computational Statistics and Data Analysis 54, 333-347

Lima Neto, E.d.A., de Carvalho, F.d.A.T. Nonlinear regression applied to interval-valued data. Pattern Anal Applic 20, 809–824 (2017). https://doi.org/10.1007/s10044-016-0538-y

Rodriguez, O. (2018). Shrinkage linear regression for symbolic interval-valued variables.Journal MODULAD 2018, vol. Modulad 45, pp.19-38

`sym.predict.symbolic_rt_cm` *Predict rt_cm model*

Description

Predict rt_cm model

Usage

```
## S3 method for class 'symbolic_rt_cm'
sym.predict(model, new.sym.data, ...)
```

Arguments

`model`	a model_rt_crm object
`new.sym.data`	new data
`...`	arguments to predict.rpart

References

Lima-Neto, E.A., De Carvalho, F.A.T., (2008). Centre and range method to fitting a linear regression model on symbolic interval data. Computational Statistics and Data Analysis52, 1500-1515

Lima-Neto, E.A., De Carvalho, F.A.T., (2010). Constrained linear regression models for symbolic interval-valued variables. Computational Statistics and Data Analysis 54, 333-347

Lima Neto, E.d.A., de Carvalho, F.d.A.T. Nonlinear regression applied to interval-valued data. Pattern Anal Applic 20, 809–824 (2017). https://doi.org/10.1007/s10044-016-0538-y

Rodriguez, O. (2018). Shrinkage linear regression for symbolic interval-valued variables.Journal MODULAD 2018, vol. Modulad 45, pp.19-38

sym.predict.symbolic_rt_crm

Predict rt_crm model

Description

Predict rt_crm model

Usage

```
## S3 method for class 'symbolic_rt_crm'
sym.predict(model, new.sym.data, ...)
```

Arguments

`model`	a model_rt_crm object
`new.sym.data`	new data
`...`	optional parameters

References

Lima-Neto, E.A., De Carvalho, F.A.T., (2008). Centre and range method to fitting a linear regression model on symbolic interval data. Computational Statistics and Data Analysis52, 1500-1515

Lima-Neto, E.A., De Carvalho, F.A.T., (2010). Constrained linear regression models for symbolic interval-valued variables. Computational Statistics and Data Analysis 54, 333-347

Lima Neto, E.d.A., de Carvalho, F.d.A.T. Nonlinear regression applied to interval-valued data. Pattern Anal Applic 20, 809–824 (2017). https://doi.org/10.1007/s10044-016-0538-y

Rodriguez, O. (2018). Shrinkage linear regression for symbolic interval-valued variables.Journal MODULAD 2018, vol. Modulad 45, pp.19-38

sym.predict.symbolic_svm_cm

Predict model_svm_cm model

Description

Predict model_svm_cm model

Usage

```
## S3 method for class 'symbolic_svm_cm'
sym.predict(model, new.sym.data, ...)
```

Arguments

`model`	model
`new.sym.data`	new data
`...`	optional parameters

References

Lima-Neto, E.A., De Carvalho, F.A.T., (2008). Centre and range method to fitting a linear regression model on symbolic interval data. Computational Statistics and Data Analysis52, 1500-1515

Lima-Neto, E.A., De Carvalho, F.A.T., (2010). Constrained linear regression models for symbolic interval-valued variables. Computational Statistics and Data Analysis 54, 333-347

Lima Neto, E.d.A., de Carvalho, F.d.A.T. Nonlinear regression applied to interval-valued data. Pattern Anal Applic 20, 809–824 (2017). https://doi.org/10.1007/s10044-016-0538-y

Rodriguez, O. (2018). Shrinkage linear regression for symbolic interval-valued variables.Journal MODULAD 2018, vol. Modulad 45, pp.19-38

`sym.predict.symbolic_svm_crm` *Predict model_svm_crm model*

Description

Predict model_svm_crm model

Usage

```
## S3 method for class 'symbolic_svm_crm'
sym.predict(model, new.sym.data, ...)
```

Arguments

`model`	model
`new.sym.data`	new data
`...`	optional parameters

References

Lima-Neto, E.A., De Carvalho, F.A.T., (2008). Centre and range method to fitting a linear regression model on symbolic interval data. Computational Statistics and Data Analysis52, 1500-1515

Lima-Neto, E.A., De Carvalho, F.A.T., (2010). Constrained linear regression models for symbolic interval-valued variables. Computational Statistics and Data Analysis 54, 333-347

Lima Neto, E.d.A., de Carvalho, F.d.A.T. Nonlinear regression applied to interval-valued data. Pattern Anal Applic 20, 809–824 (2017). https://doi.org/10.1007/s10044-016-0538-y

Rodriguez, O. (2018). Shrinkage linear regression for symbolic interval-valued variables.Journal MODULAD 2018, vol. Modulad 45, pp.19-38

sym.quantiles.PCA.plot

sym.quantiles.PCA.plot

Description

sym.quantiles.PCA.plot

Usage

```
sym.quantiles.PCA.plot(
  histogram.PCA.r,
  concept.names,
  var.names,
  Title,
  axes.x.label,
  axes.y.label,
  label.name
)
```

Arguments

`histogram.PCA.r`	A quantil matrix
`concept.names`	Concept Name
`var.names`	Variables to plot
`Title`	Plot title
`axes.x.label`	Label of axis X
`axes.y.label`	Label of axis Y
`label.name`	Concept Variable

Value

3D plot

Author(s)

Jorge Arce Garro

Examples

```
## Not run:
data("hardwoodBrito")
Hardwood.histogram<-hardwoodBrito
Hardwood.cols<-colnames(Hardwood.histogram)
Hardwood.names<-row.names(Hardwood.histogram)
```

```
 M<-length(Hardwood.cols)
 N<-length(Hardwood.names)
 BIN.Matrix<-matrix(rep(3,N*M),nrow = N)
pca.hist<-sym.histogram.pca(Hardwood.histogram,BIN.Matrix)
Hardwood.quantiles.PCA<-quantiles.RSDA(pca.hist$sym.hist.matrix.PCA,3)
label.name<-"Hard Wood"
Title<-"First Principal Plane"
axes.x.label<- "PC 1 (84.83%)"
axes.y.label<- "PC 2 (9.70%)"
concept.names<-c("ACER")
var.names<-c("PC.1","PC.2")
plot.3D.HW<-sym.quantiles.PCA.plot(Hardwood.quantiles.PCA,
                                   concept.names,
                                   var.names,
                                   Title,
                                   axes.x.label,
                                   axes.y.label,
                                   label.name)

plot.3D.HW

## End(Not run)
```

`sym.rf` *Symbolic Regression with Random Forest*

Description

Symbolic Regression with Random Forest

Usage

```
sym.rf(formula, sym.data, method = c("cm", "crm"), ntree = 500)
```

Arguments

`formula`	a formula, with a response but no interaction terms. If this a a data frame, that is taken as the model frame (see model.frame).
`sym.data`	symbolic data table
`method`	cm crm
`ntree`	Number of trees to grow. This should not be set to too small a number, to ensure that every input row gets predicted at least a few times.

References

Lima-Neto, E.A., De Carvalho, F.A.T., (2008). Centre and range method to fitting a linear regression model on symbolic interval data. Computational Statistics and Data Analysis52, 1500-1515

Lima-Neto, E.A., De Carvalho, F.A.T., (2010). Constrained linear regression models for symbolic interval-valued variables. Computational Statistics and Data Analysis 54, 333-347

Lima Neto, E.d.A., de Carvalho, F.d.A.T. Nonlinear regression applied to interval-valued data. Pattern Anal Applic 20, 809–824 (2017). https://doi.org/10.1007/s10044-016-0538-y

Rodriguez, O. (2018). Shrinkage linear regression for symbolic interval-valued variables.Journal MODULAD 2018, vol. Modulad 45, pp.19-38

`sym.rt` *Symbolic Regression Trees*

Description

Symbolic Regression Trees

Usage

```
sym.rt(
  formula,
  sym.data,
  method = c("cm", "crm"),
  minsplit = 20,
  maxdepth = 10
)
```

Arguments

`formula`	a formula, with a response but no interaction terms. If this a a data frame, that is taken as the model frame (see model.frame).
`sym.data`	a symbolic data table
`method`	cm crm
`minsplit`	the minimum number of observations that must exist in a node in order for a split to be attempted.
`maxdepth`	Set the maximum depth of any node of the final tree, with the root node counted as depth 0. Values greater than 30 rpart will give nonsense results on 32-bit machines.

References

Lima-Neto, E.A., De Carvalho, F.A.T., (2008). Centre and range method to fitting a linear regression model on symbolic interval data. Computational Statistics and Data Analysis52, 1500-1515

Lima-Neto, E.A., De Carvalho, F.A.T., (2010). Constrained linear regression models for symbolic interval-valued variables. Computational Statistics and Data Analysis 54, 333-347

Lima Neto, E.d.A., de Carvalho, F.d.A.T. Nonlinear regression applied to interval-valued data. Pattern Anal Applic 20, 809–824 (2017). https://doi.org/10.1007/s10044-016-0538-y

Rodriguez, O. (2018). Shrinkage linear regression for symbolic interval-valued variables.Journal MODULAD 2018, vol. Modulad 45, pp.19-38

`sym.scatterplot` *Symbolic Scatter Plot*

Description

This function could be use to plot two symbolic variables in a X-Y plane.

Usage

```
sym.scatterplot(sym.var.x, sym.var.y, labels = FALSE, ...)
```

Arguments

`sym.var.x`	First symbolic variable
`sym.var.y`	Second symbolic variable.
`labels`	As in R plot function.
`...`	As in R plot function.

Value

Return a graphics.

Author(s)

Oldemar Rodriguez Rojas

References

Bock H-H. and Diday E. (eds.) (2000). Analysis of Symbolic Data. Exploratory methods for extracting statistical information from complex data. Springer, Germany.

Rodriguez, O. (2000). Classification et Modeles Lineaires en Analyse des Donnees Symboliques. Ph.D. Thesis, Paris IX-Dauphine University.

See Also

sym.scatterplot3d

Examples

```
## Not run:
data(example3)
sym.data <- example3
sym.scatterplot(sym.data[, 3], sym.data[, 7], col = "blue", main = "Main Title")
sym.scatterplot(sym.data[, 1], sym.data[, 4],
  labels = TRUE, col = "blue",
  main = "Main Title"
)
```

```
sym.scatterplot(sym.data[, 2], sym.data[, 6],
  labels = TRUE,
  col = "red", main = "Main Title", lwd = 3
)

data(oils)
sym.scatterplot(oils[, 2], oils[, 3],
  labels = TRUE,
  col = "red", main = "Oils Data"
)
data(lynne1)

sym.scatterplot(lynne1[, 2], lynne1[, 1],
  labels = TRUE,
  col = "red", main = "Lynne Data"
)

## End(Not run)
```

`sym.set` *Create an symbolic_set type object*

Description

Create an symbolic_set type object

Usage

```
sym.set(x = NA)
```

Arguments

x	character vector

Value

a symbolic set

Examples

```
sym.set(factor(c("a", "b", "b", "l")))
```

sym.svm *Symbolic Support Vector Machines Regression*

Description

Symbolic Support Vector Machines Regression

Usage

```
sym.svm(
  formula,
  sym.data,
  method = c("cm", "crm"),
  scale = TRUE,
  kernel = "radial"
)
```

Arguments

`formula`	a symbolic description of the model to be fit.
`sym.data`	symbolic data.table
`method`	method
`scale`	A logical vector indicating the variables to be scaled. If scale is of length 1, the value is recycled as many times as needed. Per default, data are scaled internally (both x and y variables) to zero mean and unit variance. The center and scale values are returned and used for later predictions.
`kernel`	the kernel used in training and predicting. You might consider changing some of the following parameters, depending on the kernel type.

References

Lima-Neto, E.A., De Carvalho, F.A.T., (2008). Centre and range method to fitting a linear regression model on symbolic interval data. Computational Statistics and Data Analysis52, 1500-1515

Lima-Neto, E.A., De Carvalho, F.A.T., (2010). Constrained linear regression models for symbolic interval-valued variables. Computational Statistics and Data Analysis 54, 333-347

Lima Neto, E.d.A., de Carvalho, F.d.A.T. Nonlinear regression applied to interval-valued data. Pattern Anal Applic 20, 809–824 (2017). https://doi.org/10.1007/s10044-016-0538-y

Rodriguez, O. (2018). Shrinkage linear regression for symbolic interval-valued variables.Journal MODULAD 2018, vol. Modulad 45, pp.19-38

sym.umap *UMAP for Symbolic Data*

Description

This function applies the UMAP algorithm to a symbolic data table.

Usage

```
sym.umap(sym.data, ...)

## S3 method for class 'symbolic_tbl'
sym.umap(
  sym.data = NULL,
  config = umap::umap.defaults,
  method = c("naive", "umap-learn"),
  preserve.seed = TRUE,
  ...
)
```

Arguments

sym.data	symbolic data table
...	list of settings; values overwrite defaults from config; see documentation of umap.default for details about available settings
config	object of class umap.config
method	character, implementation. Available methods are 'naive' (an implementation written in pure R) and 'umap-learn' (requires python package 'umap-learn')
preserve.seed	logical, leave TRUE to insulate external code from randomness within the umap algorithms; set FALSE to allow randomness used in umap algorithms to alter the external random-number generator

sym.var *Symbolic Variable*

Description

This function get a symbolic variable from a symbolic data table.

Usage

```
sym.var(sym.data, number.sym.var)
```

Arguments

`sym.data`	The symbolic data table
`number.sym.var`	The number of the column for the variable (feature) that we want to get.

Value

Return a symbolic data variable with the following structure:

$N

[1] 7

$var.name

[1] 'F6'

$var.type

[1] '$I'

$obj.names

[1] 'Case1' 'Case2' 'Case3' 'Case4' 'Case5' 'Case6' 'Case7'

$var.data.vector

F6 F6.1

Case1 0.00 90.00

Case2 -90.00 98.00

Case3 65.00 90.00

Case4 45.00 89.00

Case5 20.00 40.00

Case6 5.00 8.00

Case7 3.14 6.76

Author(s)

Oldemar Rodriguez Rojas

References

Billard L. and Diday E. (2006). Symbolic data analysis: Conceptual statistics and data mining. Wiley, Chichester.

Bock H-H. and Diday E. (eds.) (2000). Analysis of Symbolic Data. Exploratory methods for extracting statistical information from complex data. Springer, Germany.

See Also

sym.obj

USCrime *Us crime classic data table*

Description

Us crime classic data table that can be used to generate symbolic data tables.

Usage

```
data(USCrime)
```

Format

An object of class `data.frame` with 1994 rows and 103 columns.

Source

http://archive.ics.uci.edu/ml/

References

HASTIE, T., TIBSHIRANI, R. and FRIEDMAN, J. (2008). The Elements of Statistical Learning: Data Mining, Inference and Prediction. New York: Springer.

Examples

```
## Not run:
data(USCrime)
us.crime <- USCrime
dim(us.crime)
head(us.crime)
summary(us.crime)
names(us.crime)
nrow(us.crime)
```

```
result <- classic.to.sym(us.crime,
  concept = "state",
  variables = c(NumInShelters, NumImmig),
  variables.types = c(
    NumInShelters = type.histogram(),
    NumImmig = type.histogram()
  )
)
result

## End(Not run)
```

---

uscrime_int *Us crime interval data table.*

Description

Us crime classic data table genetated from uscrime data.

Usage

```
data(uscrime_int)
```

Format

An object of class `symbolic_tbl` (inherits from `tbl_df`, `tbl`, `data.frame`) with 46 rows and 102 columns.

References

Rodriguez O. (2013). A generalization of Centre and Range method for fitting a linear regression model to symbolic interval data using Ridge Regression, Lasso and Elastic Net methods. The IFCS2013 conference of the International Federation of Classification Societies, Tilburg University Holland.

Examples

```
data(uscrime_int)
car.data <- uscrime_int
res.cm.lasso <- sym.glm(
  sym.data = car.data, response = 102, method = "cm", alpha = 1,
  nfolds = 10, grouped = TRUE
)
plot(res.cm.lasso)
plot(res.cm.lasso$glmnet.fit, "norm", label = TRUE)
plot(res.cm.lasso$glmnet.fit, "lambda", label = TRUE)

pred.cm.lasso <- sym.predict(res.cm.lasso, response = 102, car.data)
RMSE.L(car.data$ViolentCrimesPerPop, pred.cm.lasso)
RMSE.U(car.data$ViolentCrimesPerPop, pred.cm.lasso)
```

```
R2.L(car.data$ViolentCrimesPerPop, pred.cm.lasso)
R2.U(car.data$ViolentCrimesPerPop, pred.cm.lasso)
deter.coefficient(car.data$ViolentCrimesPerPop, pred.cm.lasso)
```

uscrime_intv2 *Us crime interval data table.*

Description

Us crime classic data table genetated from uscrime data.

Usage

```
data(uscrime_int)
```

Format

An object of class `symbolic_tbl` (inherits from `tbl_df`, `tbl`, `data.frame`) with 46 rows and 102 columns.

References

Rodriguez O. (2013). A generalization of Centre and Range method for fitting a linear regression model to symbolic interval data using Ridge Regression, Lasso and Elastic Net methods. The IFCS2013 conference of the International Federation of Classification Societies, Tilburg University Holland.

var *Symbolic Variance*

Description

Compute the symbolic variance.

Usage

```
var(x, ...)

## Default S3 method:
var(x, y = NULL, na.rm = FALSE, use, ...)

## S3 method for class 'symbolic_interval'
var(x, method = c("centers", "interval", "billard"), na.rm = FALSE, ...)

## S3 method for class 'symbolic_tbl'
var(x, ...)
```

Arguments

`x`	A symbolic interval.
`...`	As in R median function.
`y`	NULL (default) or a vector, matrix or data frame with compatible dimensions to x. The default is equivalent to y = x (but more efficient).
`na.rm`	logical. Should missing values be removed?
`use`	an optional character string giving a method for computing covariances in the presence of missing values. This must be (an abbreviation of) one of the strings 'everything', 'all.obs', 'complete.obs', 'na.or.complete', or 'pairwise.complete.obs'.
`method`	The method to be use.

Author(s)

Oldemar Rodriguez Rojas

References

Billard L. and Diday E. (2006). Symbolic data analysis: Conceptual statistics and data mining. Wiley, Chichester.

Rodriguez, O. (2000). Classification et Modeles Lineaires en Analyse des Donnees Symboliques. Ph.D. Thesis, Paris IX-Dauphine University.

`variance.princ.curve` *Variance of the principal curve*

Description

Variance of the principal curve

Usage

```
variance.princ.curve(data,curve)
```

Arguments

`data`	Classic data table.
`curve`	The principal curve.

Value

The variance of the principal curve.

Author(s)

Jorge Arce.

References

Arce J. and Rodriguez O. (2015) 'Principal Curves and Surfaces to Interval Valued Variables'. The 5th Workshop on Symbolic Data Analysis, SDA2015, Orleans, France, November.

Hastie,T. (1984). Principal Curves and Surface. Ph.D Thesis Stanford University.

Hastie,T. & Weingessel,A. (2014). princurve - Fits a Principal Curve in Arbitrary Dimension.R package version 1.1–12 http://cran.r-project.org/web/packages/princurve/index.html.

Hastie,T. & Stuetzle, W. (1989). Principal Curves. Journal of the American Statistical Association, Vol. 84-406, 502–516.

Hastie, T., Tibshirani, R. & Friedman, J. (2008). The Elements of Statistical Learning; Data Mining, Inference and Prediction. Springer, New York.

See Also

sym.interval.pc

`vertex.interval` *Vertex of the intervals*

Description

Vertex of the intervals

Usage

```
vertex.interval(sym.data)
```

Arguments

`sym.data`	Symbolic interval data table.

Value

Vertices of the intervals.

Author(s)

Jorge Arce.

References

Arce J. and Rodriguez O. (2015) 'Principal Curves and Surfaces to Interval Valued Variables'. The 5th Workshop on Symbolic Data Analysis, SDA2015, Orleans, France, November.

Hastie,T. (1984). Principal Curves and Surface. Ph.D Thesis Stanford University.

Hastie,T. & Weingessel,A. (2014). princurve - Fits a Principal Curve in Arbitrary Dimension.R package version 1.1–12 http://cran.r-project.org/web/packages/princurve/index.html.

Hastie,T. & Stuetzle, W. (1989). Principal Curves. Journal of the American Statistical Association, Vol. 84-406, 502–516.

Hastie, T., Tibshirani, R. & Friedman, J. (2008). The Elements of Statistical Learning; Data Mining, Inference and Prediction. Springer, New York.

See Also

sym.interval.pc

`VeterinaryData` *Symbolic interval data example*

Description

Symbolic data matrix with all the variables of interval type.

Usage

```
data(VeterinaryData)
```

Format

$I Height Height $I Weight Weight

1 $I 120.0 180.0 $I 222.2 354.0

2 $I 158.0 160.0 $I 322.0 355.0

3 $I 175.0 185.0 $I 117.2 152.0

4 $I 37.9 62.9 $I 22.2 35.0

5 $I 25.8 39.6 $I 15.0 36.2

6 $I 22.8 58.6 $I 15.0 51.8

7 $I 22.0 45.0 $I 0.8 11.0

8 $I 18.0 53.0 $I 0.4 2.5

9 $I 40.3 55.8 $I 2.1 4.5

10 $I 38.4 72.4 $I 2.5 6.1

References

Billard L. and Diday E. (2006). Symbolic data analysis: Conceptual statistics and data mining. Wiley, Chichester.

Examples

```
data(VeterinaryData)
VeterinaryData
```

`weighted.center.Hist.RSDA` *weighted.center.Hist.RSDA*

Description

weighted.center.Hist.RSDA

Usage

```
weighted.center.Hist.RSDA(sym.histogram)
```

Arguments

`sym.histogram` A Histogram matrix

Value

Matrix of Weighted Centers

Author(s)

Jorge Arce Garro

Examples

```
## Not run:
data(hardwoodBrito)
weighted.center.Hist.RSDA(hardwoodBrito)

## End(Not run)
```

write.sym.table — *Write Symbolic Data Table*

Description

This function write (save) a symbolic data table from a CSV data file.

Usage

```
write.sym.table(sym.data, file, sep, dec, row.names = NULL, col.names = NULL)
```

Arguments

`sym.data`	Symbolic data table
`file`	The name of the CSV file.
`sep`	As in R function read.table
`dec`	As in R function read.table
`row.names`	As in R function read.table
`col.names`	As in R function read.table

Value

Write in CSV file the symbolic data table.

Author(s)

Oldemar Rodriguez Rojas

References

Bock H-H. and Diday E. (eds.) (2000). Analysis of Symbolic Data. Exploratory methods for extracting statistical information from complex data. Springer, Germany.

See Also

read.sym.table

Examples

```
## Not run:
data(example1)
write.sym.table(example1, file = "temp4.csv", sep = "|",
                dec = ".", row.names = TRUE, col.names = TRUE)
ex1 <- read.sym.table("temp4.csv", header = TRUE,
                       sep = "|", dec = ".", row.names = 1)

## End(Not run)
```

$.symbolic_histogram *$ operator for histograms*

Description

$ operator for histograms

Usage

```
## S3 method for class 'symbolic_histogram'
x$name
```

Arguments

x	.....
name	...

$.symbolic_modal *$ operator for modals*

Description

$ operator for modals

Usage

```
## S3 method for class 'symbolic_modal'
x$name = c("cats", "props", "counts")
```

Arguments

x	.....
name	...

`$.symbolic_set` *$ operator for set*

Description

$ operator for set

Usage

```
## S3 method for class 'symbolic_set'
x$name = c("levels", "values")
```

Arguments

x	.....
name	...

Index