Package: ordinalLBM (via r-universe)

September 9, 2024

Title Co-Clustering of Ordinal Data via Latent Continuous Random Variables

Version 1.0

Author Marco Corneli, Charles Bouveyron and Pierre Latouche

Maintainer Marco Corneli <marcogenni@gmail.com>

Description It implements functions for simulation and estimation of the ordinal latent block model (OLBM), as described in Corneli, Bouveyron and Latouche (2019).

Imports reshape2, RColorBrewer

Depends R (>= 3.4.0)

License GPL (>= 2)

Encoding UTF-8

LazyData true

RoxygenNote 6.1.0

NeedsCompilation no

Date/Publication 2019-01-30 23:20:04 UTC

Repository https://marcogenni.r-universe.dev

RemoteUrl https://github.com/cran/ordinalLBM

RemoteRef HEAD

RemoteSha 42321c99b68be5470e02586e9a29f665fbde451c

Contents

olbm 2
olbm_dat 3
plot.olbm 4
simu.olbm 4

Index **6**

olbm *Fitting OLBM to the data*

Description

It estimates the OLBM model parameters as well as the most likely posterior cluster assignments by maximum likelihood.

Usage

```
olbm(Y, Q, L, init = "kmeans", eps = 1e-04, it_max = 500,
  verbose = TRUE)
```

Arguments

Y	An M x P ordinal matrix, containing ordinal entries from 1 to K. Missing data are coded as zeros.
Q	The number of row clusters.
L	The number of column clusters.
init	A string specifying the initialisation type. It can be "kmeans" (the default) or "random" for a single random initialisation.
eps	When the difference between two consecutive vaules of the log-likelihood is smaller than eps, the M-EM algorithms will stop.
it_max	The maximum number of iterations that the M-EM algorithms will perform (although the minimum tolerance eps is not reached).
verbose	A boolean specifying whether extended information should be displayed or not (TRUE by default).

Value

It returns an S3 object of class "olbm" containing

estR	the estimated row cluster memberships.
estC	the estimated column cluster memberships.
likeli	the final value of the log-likelihood.
icl	the value of the ICL criterion.
Pi	the Q x L estimated connectivity matrix.
mu	a Q x L matrix containing the estimated means of the latent Gaussian distributions.
sd	a Q x L matrix containing the estimated standard deviations of the latent Gaussian distributions.
eta	a Q x L x K array whose entry (q,l,k) is the estimated probability that one user in the q-th row cluster assign the score k to one product in the l-th column cluster.

`rho`	the estimated row cluster proportions.
`delta`	the estimated column cluster proportions.
`initR`	the initial row cluster assignments provided to the C-EM algorithm.
`initC`	the initial column cluter assignments provided to the C-EM algorigthm.
`Y`	the input ordinal matrix Y.
`thresholds`	the values (1.5, 2.5, ... , K-0.5) of the thresholds, defined inside the function olbm.

References

Corneli M.,Bouveyron C. and Latouche P. (2019) *Co-Clustering of ordinal data via latent continuous random variables and a classification EM algorithm.* (https://hal.archives-ouvertes.fr/hal-01978174)

Examples

```
data(olbm_dat)
res <- olbm(olbm_dat$Y, Q=3, L=2)
```

`olbm_dat` *OLBM simulated data*

Description

It is a list containing i) an ordinal toy data matrix simulated acccording to OLBM and ii) the row/column cluster assignments. To see how the data are simulated, you can type "?simu.olbm" in the R console and look at "Examples".

Usage

```
data(olbm_dat)
```

Format

A list containing three items.

Y : an ordinal data matrix simulated according to OLBM.

Rclus : the actual row cluster assignments.

Cclust : the actual column cluster assignments.

plot.olbm *Plot OLBM*

Description

It plots the re-organized incidence matrix and/or the estimated Gussian densities.

Usage

```
## S3 method for class 'olbm'
plot(x, type = "hist", ...)
```

Arguments

x	The "olbm" object output of the function olbm.
type	A string specifying the type of plot to be produced. The currently supported values are "hist" and "incidence".
...	Additional parameters to pass to sub-functions.

Examples

```
data(olbm_dat)
res <- olbm(olbm_dat$Y, Q=3, L=2)
plot(res, "hist")
plot(res, "incidence")
```

simu.olbm *Simulate OLBM data*

Description

It simulates an ordinal data matrix according to OLBM.

Usage

```
simu.olbm(M, P, Pi, rho, delta, mu, sd, thresh)
```

Arguments

M	The number of rows of the ordinal matrix Y.
P	The number of columns of the ordinal matrix Y.
Pi	A Q x L connectivity matrix to manage missing data (coded az zeros in Y).
rho	A vector of length Q, containing multinomial probabilities for row cluster assignments.

delta	A vector of length L, containing multinomial probabilities for column cluster assignments.
mu	A Q x L matrix containing the means of the latent Gaussian distributions.
sd	A Q x L matrix containing the standard deviations of the latent Gaussian distributions.
thresh	A K+1 vector containing the sorted tresholds used to simulate the ordinal entries in Y, where K is the number of ordinal modalities. The first entry in tresh must be -Inf, the last entry +Inf.

Value

It returns a list containing:

Y	An M x P matrix. The observed ordinal entries are integers between 1 and K. Missing data are coded as zeros.
Rclus	A vector of length M containing the row cluster memberships.
Cclus	A vector of length P containing the column cluster memberships.

References

Corneli M.,Bouveyron C. and Latouche P. (2019) *Co-Clustering of ordinal data via latent continuous random variables and a classification EM algorithm.* (https://hal.archives-ouvertes.fr/hal-01978174)

Examples

```
M <- 150
P <- 100
Q <- 3
L <- 2

## connectivity matrix
Pi <- matrix(.7, nrow = Q, ncol = L)
Pi[1,1] <- Pi[2,2] <- Pi[3,2] <- .5

## cluster memberships proportions
rho <- c(1/3, 1/3 ,1/3)
delta <- c(1/2, 1/2)

# Thresholds
thresh <- c(-Inf, 2.37, 2.67, 3.18, 4.33, Inf)     # K = 5

## Gaussian parameters
mu <- matrix(c(0, 3.4, 2.6, 0, 2.6, 3.4), nrow = Q, ncol = L)
sd <- matrix(c(1.2,1.4,1.0,1.2,1.4,1.0), nrow = Q, ncol = L)

## Data simulation
dat <- simu.olbm(M, P, Pi, rho, delta, mu, sd, thresh)
```

Index

olbm, 2
olbm_dat, 3

plot.olbm, 4

simu.olbm, 4