Package 'cobiclust'

October 15, 2018

Type Package

Title Biclustering via Latent Block Model Adapted to Overdispersed Count Data

Version 0.1.0

Author Julie Aubert

Maintainer Julie Aubert <julie.aubert@agroparistech.fr>

Description Implementation of a probabilistic method for biclustering adapted to overdispersed count data. It is a Gamma-Poisson Latent Block Model. It also implements two selection criteria in order to select the number of biclusters.

Imports cluster

License GPL-2

Encoding UTF-8

LazyData true

RoxygenNote 6.0.1

NeedsCompilation no

Repository CRAN

Date/Publication 2018-10-15 18:50:07 UTC

R topics documented:

`cobiclust` *Perform a biclustering adapted to overdispersed count data.*

Description

Perform a biclustering adapted to overdispersed count data.

Usage

```
cobiclust(x, K = 2, G = 3, nu_j = NULL, a = NULL, akg = FALSE,
  cvg_lim = 1e-05, nbiter = 5000)
```

Arguments

`x`	the input matrix of observed data.
`K`	an integer specifying the number of groups in rows.
`G`	an integer specifying the number of groups in columns.
`nu_j`	a vector of . The length is equal to the number of colums.
`a`	an numeric.
`akg`	a logical variable indicating whether to use a common dispersion parameter (akg = FALSE) or a dispersion parameter per cocluster (akg = TRUE).
`cvg_lim`	a number specifying the threshold used for convergence criterion (cvg_lim = 1e-05 by default).
`nbiter`	the maximal number of iterations for the global loop of variational EM algorithm (nbiter = 5000 by default).

Value

An object of class cobiclustering

See Also

`cobiclustering` for the cobiclustering class.

Examples

```
npc <- c(50, 40) # nodes per class
KG <- c(2, 3) # classes
nm <- npc * KG # nodes
Z <- diag( KG[1]) %x% matrix(1, npc[1], 1)
W <- diag(KG[2]) %x% matrix(1, npc[2], 1)
L <- 70 * matrix( runif( KG[1] * KG[2]), KG[1], KG[2])
M_in_expectation <- Z %*% L %*% t(W)
size <- 50
M<-matrix(
 rnbinom(
```

```
    n = length(as.vector(M_in_expectation)),
    mu = as.vector(M_in_expectation), size = size)
  , nm[1], nm[2])
rownames(M) <- paste("OTU", 1:nrow(M), sep = "_")
colnames(M) <- paste("S", 1:ncol(M), sep = "_")
res <- cobiclust(M, K = 2, G = 3, nu_j = rep(1,120), a = 1/size, cvg_lim = 1e-5)
```

`selection_criteria` *Calculate selection criteria.*

Description

Calculate selection criteria.

Usage

```
selection_criteria(x, K, G)
```

Arguments

`x`	The output of the cobiclust function.
`K`	The number of groups in rows.
`G`	The number of groups in columns.

Value

A dataframe with 7 columns.

`vICL` the vICL selection criterion.

`BIC` the BIC selection criterion.

`penKG` the value of the BIC penalty.

`lb` the value of the lower bound of the log-likelihood.

`entZW` the value of the entropy of the latent variables Z and W.

`K` the number of groups in rows.

`G` the number of groups in columns.

Index