Package 'MatManlyMix'

January 5, 2018

Version 1.1.1

Date 2017-12-25

Title Matrix Clustering with Gaussian and Manly Mixture Models

Depends R (>= 3.0.0)

LazyLoad yes

LazyData no

Description Matrix clustering with finite mixture models.

License GPL (>= 2)

Author Xuwen Zhu [aut, cre],
Volodymyr Melnykov [aut],
Michael Hutt [ctb, cph] (NM optimization in c),
Stephen Moshier [ctb, cph] (eigen calculations in c),
Rouben Rostamian [ctb, cph] (memory allocation in c)

Maintainer Xuwen Zhu <xuwen.zhu@louisville.edu>

NeedsCompilation yes

Repository CRAN

Date/Publication 2018-01-05 08:20:25 UTC

R topics documented:

MatManlyMix-package   *Finite mixture modeling and model-based clustering of matrices based*
                       on matrix Gaussian mixture and matrix Manly mixture models.

Description

The utility of this package is the clustering of random matrices. Finite mixture modeling and model-based clustering based on matrix Gaussian mixtures and matrix Manly mixtures are employed.

Details

Package:	MatManlyMix
Type:	Package
Version:	0.1.1
Date:	2017-02-09
License:	GPL (>= 2)
LazyLoad:	no

Function 'MatManly.init' runs the initialization for the EM algorithm.

Function 'MatManly.EM' runs the EM algorithm for matrix-variate mixtures to cluster matrices.

Author(s)

Xuwen Zhu and Volodymyr Melnykov.

Maintainer: Xuwen Zhu <xuwen.zhu@louisville.edu>

---

crime   *Crime data*

Description

Data collected by FBI's Uniform Crime on the violent and property crimes of 236 cities.

Usage

```
data(crime)
```

Format

A list of 3 objects: Y, department and state. Y represents the crime rate data array from 236 cities. Department is the police department names and state represents the states where each city is located at. Y is of dimensionality 10 x 13 x 236 with 236 crime rates on the following 10 variables from year 2000 through 2012.

Population Population of each city;

Violent Crime rate Total number of violent crimes;

Murder and non-negligent manslaughter rate Number of murders;

Forcible rape rate Number of rape crimes;

Robbery rate Number of robberies;

Aggravated assault rate Number of assaults;

Property crime rate Total number of property crimes;

Burglary rate Number of burglary crimes;

Larceny-theft rate Number of theft crimes;

Motor vehicle theft rate Number of vehicle theft crimes;

Details

The data have been made publicly available by FBI's Uniform Crime Reports.

Examples

```
data(crime)
```

IMDb *IMDb data*

Description

Data collected from IMDb.com on the ratings of 105 popular comedy movies.

Usage

```
data(IMDb)
```

Format

A list of 2 objects: Y and name, where Y represents the data array of ratings and name represents the comedy movie names. Y is the of dimensionality 2 x 4 x 105 with ratings on 105 movies from female and male by age groups 0-18, 18-29, 30-44, 45+.

Details

The data are publicly available on www.IMDb.com.

Examples

```
data(IMDb)
```

MatManly.EM   *EM algorithm for matrix clustering*

Description

Runs the EM algorithm for matrix clustering

Usage

```
MatManly.EM(Y, X = NULL, initial = NULL, id = NULL, la = NULL, nu = NULL, tau = NULL,
Mu = NULL, beta = NULL, Sigma = NULL, Psi = NULL, Mu.type = 0, Psi.type = 0,
tol = 1e-05, max.iter = 1000, size.control = 0, silent = TRUE)
```

Arguments

Y	dataset of random matrices (p x T x n), n random matrices of dimensionality (p x T)
X	dataset of explanatory variables (T x q x n), q explanatory variables for modeling Y
initial	initialization parameters provided by function MatManly.init()
id	initial membership vector
la	initial transformation parameters (K x p)
nu	initial transformation parameters (K x T)
tau	initial mixing proportions (length K)
Mu	initial mean matrices (p x T x K)
beta	initial coefficient matrices (q x p x K)
Sigma	initial array of sigma (p x p x K)
Psi	initial array of psi (T x T x K)
Mu.type	mean structure: 0-unrestricted, 1-additive
Psi.type	covariance structure of the Psi matrix: 0-unrestricted, 1-AR1
tol	tolerance level
max.iter	maximum number of iterations
size.control	minimum size of clusters allowed for controlling spurious solutions
silent	whether to produce output of steps or not

Details

Runs the EM algorithm for modeling and clustering matrices for a provided dataset. Both matrix Gaussian mixture and matrix Manly mixture with given explanatory variables (data matrix X) or without explanatory variables (X is null) can be employed. A user has three options to initialize the EM algorithm. The user can use the MatManly.init() function to get initial parameters and input them as 'initial'. The second choice is to specify either initial id vector 'id' and transformation

parameters 'la'. The third option is to input initial mode parameters 'la', 'tau', 'Mu', and 'Sigma' and 'Psi'. In the case when transformation parameters are not provided, the function runs the EM algorithm without any transformations, i.e., it is equivalent to the EM algorithm for a Gaussian mixture model. If some transformation parameters have to be excluded from the consideration, in the corresponding positions of matrix 'la', the user has to specify value 0. A user also has three options to specify the covariance structure of the 'Psi' matrices, including unrestricted case, spherical matrices and autoregressive structured matrices. Notation: n - sample size, p x T - dimensionality of the random matrices, K - number of mixture components.

Value

la	matrix of the skewness parameters (K x p)
nu	matrix of the skewness parameters (K x T)
tau	vector of mixing proportions (length K)
Mu	matrix of the estimated mean matrices (p x T x K)
beta	matrix of the coefficient parameters (q x p x K)
Sigma	array of the estimated sigma (p x p x K)
Psi	array of the estimated psi (T x T x K)
Mu.type	mean structure: 0-unrestricted, 1-additive
Psi.type	covariance structure of the Psi matrix: 0-unrestricted, 1-AR1
gamma	matrix of the estimated posterior probabilities (n x K)
id	estimated membership vector (length n)
ll	log likelihood value
bic	Bayesian Information Criterion
iter	number of EM iterations run
flag	convergence flag (0 - success, 1 - failure)

Examples

```
set.seed(123)

data(crime)


Y <- crime$Y[c(2,7),,] / 1000

p <- dim(Y)[1]
T <- dim(Y)[2]
n <- dim(Y)[3]
K <-  2

#init <- MatManly.init(Y, K = K, la = matrix(0.1, K, p), nu = matrix(0.1, K, T))
#M1 <- MatManly.EM(Y, initial = init, max.iter = 1000)
```

MatManly.init	*Initialization for the EM algorithm for matrix clustering*

Description

Runs the initialization for the EM algorithm for matrix clustering

Usage

```
MatManly.init(Y, X = NULL, K, la = NULL, nu = NULL, Mu.type = 0,
Psi.type = 0, n.start = 10, tol = 1e-05)
```

Arguments

Y	dataset of random matrices (p x T x n), n random matrices of dimensionality (p x T)
X	dataset of explanatory variables (T x q x n), q explanatory variables for modeling Y
K	number of clusters
la	initial transformation parameters (K x p)
nu	initial transformation parameters (K x T)
Mu.type	mean structure: 0-unrestricted, 1-additive
Psi.type	covariance structure of the Psi matrix: 0-unrestricted, 1-AR1
n.start	initial random starts
tol	tolerance level

Details

Random starts are used to obtain different starting values. The number of clusters, the skewness parameters, and number of random starts need to be specified. In the case when transformation parameters are not provided, the function runs the EM algorithm without any transformations, i.e., it is equivalent to the EM algorithm for a matrix Gaussian mixture. Notation: n - sample size, p x T - dimensionality of the random matrices, K - number of mixture components.

Satellite	*Satellite data*

Description

Data publicly available at the University of California - Irvine machine learning repository (http://archive.ics.uci.edu/ml), was originally obtained by NASA.

Usage

```
data(IMDb)
```

Format

A list of 2 objects: Y and id, where Y represents the data array of spectral values and id represents the true id of three classes: Soil with vegetation stubble, damp grey soil, and grey soil. Y is the of dimensionality 4 x 9 x 845.

Details

The data are publicly available on http://archive.ics.uci.edu/ml.

Examples

```
data(Satellite)
```

Index