Package 'qclust'

April 16, 2015

Type Package

Title Robust Estimation of Gaussian Mixture Models

Version 1.0

Date 2015-04-15

Author Yichen Qin, Carey E. Priebe

Maintainer Yichen Qin <qinyn@ucmail.uc.edu>

Description Robust estimation of Gaussian mixture models fitted by modified EM algorithm, robust clustering and classification.

Depends mclust, mvtnorm

URL http://homepages.uc.edu/~qinyn/qclust/

License GPL-2

NeedsCompilation no

Repository CRAN

Date/Publication 2015-04-16 19:53:55

R topics documented:

qclust-package *Robust Estimation of Gaussian Mixture Models for Model-Based Clustering*

Description

Finite Gaussian mixture models fitted by the robust EM algorithm, and robust model-based clustering for multidimensional data.

Details

Package:	qclust
Type:	Package
Version:	1.0
Date:	2015-04-15
License:	GPL-2
URL:	http://homepages.uc.edu/~qinyn/qclust/

The main function in this package is Qclust(). It provides a robust estimation of Gaussian mixture models and also generates a robust model-based clustering.

Author(s)

Yichen Qin, Carey E. Priebe
Maintainer: Yichen Qin <qinyn@ucmail.uc.edu>
URL: http://homepages.uc.edu/~qinyn/qclust/

dmm_fn *Component Densities of Gaussian Mixture Models with Associate Proportions*

Description

Component densities of Gaussian mixture models multiplied by its associated proportion for each data points.

Usage

```
dmm_fn(x, means, vars, pis)
```

Arguments

x	A matrix of n by p. n is the sample size. p is the dimension.
means	Means of the Gaussian mixture model.
vars	Variances of the Gaussian mixture model.
pis	Proportions of the Gaussian mixture model.

Value

A matrix of n by m. n is the sample size. m is the complexity. Each column represents the data points' densities (multiplied by its associated proportion) for one component.

Author(s)

Yichen Qin, Carey E. Priebe Maintainer: Yichen Qin <qinyn@ucmail.uc.edu>

---

dmm_noweights_fn *Component Densities of Gaussian Mixture Models*

Description

Component densities of Gaussian mixture models for each data points.

Usage

```
dmm_noweights_fn(x, means, vars, pis)
```

Arguments

x	A matrix of n by p. n is the sample size. p is the dimension.
means	Means of the Gaussian mixture model.
vars	Variances of the Gaussian mixture model.
pis	Proportions of the Gaussian mixture model.

Value

A matrix of n by m. n is the sample size. m is the complexity. Each column represents the data points' densities for one component.

Author(s)

Yichen Qin, Carey E. Priebe Maintainer: Yichen Qin <qinyn@ucmail.uc.edu>

Lq_fn *Lq Function*

Description

Defined as Lq(x)=(x^1-q-1)/(1-q), where q is a tuning parameter.

Usage

```
Lq_fn(x, q)
```

Arguments

x	A vector or a scalar.
q	A tuning parameter, 0<q<=1.

Value

The function returns (x^1-q-1)/(1-q).

Author(s)

Yichen Qin, Carey E. Priebe Maintainer: Yichen Qin <qinyn@ucmail.uc.edu>

Qclust *Robust Estimation of Gaussian Mixture Models*

Description

Robust estimation of Gaussian mixture models and robust model-based clustering.

Usage

```
Qclust(d, K = NULL, modelNames = NULL, q)
```

Arguments

d	A data matrix, n by p. n is sample size, p is the dimension.
K	The complexity of the mixture model, an integer.
modelNames	A string indicate the type of "models". The notation is consistent with the package mclust.
q	A tuning parameter, 0<q<=1. q<1 provides robust estimation.

Value

The function returns a list which contains:

`parameters` Estimated parameters of Gaussian mixture models

Author(s)

Yichen Qin, Carey E. Priebe
Maintainer: Yichen Qin <qinyn@ucmail.uc.edu>
URL: http://homepages.uc.edu/~qinyn/qclust/

References

Ferrari, D. and Yang, Y., (2010). Maximum Lq-likelihood estimation, Annals of Statistics.
Fraley, C., Raftery, A. E., Murphy, T. B., and Scrucca, L., (2012). mclust Version 4 for R: Normal Mixture Modeling for Model-Based Clustering, Classification, and Density Estimation. Technical Report No. 597, Department of Statistics, University of Washington.
Fraley, C., and Raftery, A.E., (2002). Model-Based Clustering, Discriminant Analysis, and Density Estimation. Journal of the American Statistical Association, 97, 611-631.
McLachlan, G.J., and Peel, D., (2000). Finite Mixture Models, Wiley, ISBN 0-471-00626-2.

Examples

```
n = 200
set.seed(12345)
true_para=list()
true_para$pro=rep(1/3,3)
true_para$mean=matrix(c(-6,1.5,0,0,6,1.5),2,3)
true_para$variance$sigma=array(c(5,4,4,5,5,-4,-4,5,1.56,0,0,1.56),dim=c(2,2,3))
G=ncol(true_para$mean)
z = sample(1:G,n,true_para$pro,replace=TRUE)
z = sort(z)
X=matrix(NA,n,nrow(true_para$mean))
for (i in 1:G)
{
  X[z==i,]=rmvnorm(sum(z==i),true_para$mean[,i],true_para$variance$sigma[,,i])
}
plot(X,pch=20)
qfit=Qclust(X,K=3,modelNames="VVV",q=0.9)
plot(qfit,what="classification")
```

---

qclust_w_initialvalues

Estimate Gaussian Mixture Models with Initial Values

Description

Robust estimation of Gaussian mixture models with given initial values given as arguement.

Usage

```
qclust_w_initialvalues(d, means_init, vars_init, pis_init, q, tol = 1e-05)
```

Arguments

d	A data matrix, n by p. n is sample size, p is the dimension.
means_init	Initial values for means.
vars_init	Initial values for variances.
pis_init	Initial values for proportions.
q	A tuning parameter, 0<q<=1.
tol	Tolerance level for the EM algorithm to converge.

Value

This function returns a list which contains:

means	Estimated means
vars	Estimated variances
pis	Estimated proportions

Author(s)

Yichen Qin, Carey E. Priebe Maintainer: Yichen Qin <qinyn@ucmail.uc.edu>

Index