Package 'sphereclust'

June 1, 2026

Type Package

Title Model Based Clustering for Spherical Data Using Elliptically Symmetric Distributions

Version 1.0

Date 2026-05-28

Author Michail Tsagris [aut, cre], Theodoros Perdikis [ctb]

Maintainer Michail Tsagris <mtsagris@uoc.gr>

Depends R (>= 4.0)

Imports Directional, graphics, grDevices, mixture, parallel, rangen, rgl, Rfast, stats

Suggests Rfast2

Description Model based clustering with spherical data using mixtures of elliptically symmetric distributions, namely mixtures of spherical elliptically symmetric projected Cauchy (SESPC) or mixtures of elliptically symmetric angular Gaussian (ESAG) distributions. The relevant paper is: Perdikis T., Alharbi N. and Tsagris M. (2026). <doi:10.48550/arXiv.2605.27496>.

License GPL (>= 2)

NeedsCompilation no

Repository CRAN

Date/Publication 2026-06-01 08:50:02 UTC

Contents

sphereclust-package *Model Based Clustering for Spherical Data Using Elliptically Symmetric Distributions*

Description

Model based clustering for spherical data using mixtures of spherical elliptically symmetric projected Cauchy (SESPC) or mixtures of elliptically symmetric angular Gaussian (ESAG) distributions

Details

Package:	sphereclust
Type:	Package
Version:	1.0
Date:	2026-05-28
License:	GPL-2

Maintainers

Michail Tsagris <mtsagris@uoc.gr>

Author(s)

Michail Tsagris <mtsagris@uoc.gr> and and Theodoros Perdikis <theoperdikis92@gmail.com>.

References

Perdikis T., Alharbi N. and Tsagris M. (2026). Model–based clustering for spherical and hyper–spherical data using elliptically symmetric distributions.

https://arxiv.org/abs/2605.27496

BIC and ICL for the model based clustering with elliptically symmetric distributions

BIC and ICL to choose the number of components in a model based clustering with elliptically symmetric distributions

Description

BIC and ICL to choose the number of components in a model based clustering with elliptically symmetric distributions

Usage

```
bic.mixsespc(x, G = 5, tol = 1e-4, ncores = 1)
bic.mixesag(x, G = 5, tol = 1e-4, ncores = 1)
```

Arguments

x	A matrix containing directional data.
G	The maximum number of clusters to be tested. Default value is 5.
tol	The tolerance value to terminate the EM algorithm.
ncores	The number of cores in case you want to run the search for the optimal number of clusters in parallel.

Details

The function computes the BIC and ICL to decide on the optimal number of clusters when using mixtures of SESPC or mixtures of ESAG distributions.

Value

A plot of the ICL values and a list including:

bic	The BIC values for all the models tested.
icl	The ICL values for all the models tested.
runtime	The run time of the algorithm. A numeric vector. The first element is the user time, the second element is the system time and the third element is the elapsed time.

Author(s)

Michail Tsagris.

R implementation and documentation: Michail Tsagris <mtsagris@uoc.gr> and Theodoros Perdikis <theoperdikis92@gmail.com>.

References

Perdikis T., Alharbi N. and Tsagris M. (2026). Model–based clustering for spherical and hyper–spherical data using elliptically symmetric distributions.

https://arxiv.org/abs/2605.27496

See Also

mixsespc.mle, rmixsespc

Examples

```
x <- as.matrix( iris[, 1:3] )
x <- x / sqrt( rowSums(x^2) )
bic.mixsespc(x, G = 3)
```

`Contour plot (on the sphere) of a mixture of elliptically symmetric distributions`

Contour plot (on the sphere) of a mixture of elliptically symmetric distributions

Description

The contour plot (on the sphere) of a mixture of elliptically symmetric distributions.

Usage

```
mixsespc.contour(probs, mu, theta, bgcol = "snow", dat = NULL, col = NULL,
lat = 50, long = 50)
mixesag.contour(probs, mu, gam, bgcol = "snow", dat = NULL, col = NULL,
lat = 50, long = 50)
```

Arguments

`probs`	This is avector with the mixing probability of each group.
`mu`	A matrix with the mean direction of each group.
`theta`	A matrix with the θ parameters of the SESPC distribution of each group.
`gam`	A matrix with the γ parameters of the ESAG distribution of each group.
`bgcol`	The color of the surface of the sphere.
`dat`	If you have you want to plot supply them here. This has to be a numerical matrix with three columns, i.e. unit vectors.
`col`	If you supplied data then choose the color of the points. If you did not choose a color, the points will appear in red.
`lat`	A positive number determing the range of degrees to move left and right from the latitude center. See the example to better understand this argument.
`long`	A positive number determing the range of degrees to move up and down from the longitude center. See the example to better understand this argument.

Details

The goal of this function is for the user to see how the mixtures of SESPC or ESAG distributions look like.

Value

A plot containing the contours of the mixture distribution.

Author(s)

Michail Tsagris.

R implementation and documentation: Michail Tsagris <mtsagris@uoc.gr>.

References

Perdikis T., Alharbi N. and Tsagris M. (2026). Model–based clustering for spherical and hyper–spherical data using elliptically symmetric distributions.

https://arxiv.org/abs/2605.27496

See Also

`dmixsespc`, `mixsespc.mle`

Examples

```
probs <- c(0.5, 0.5)
mu <- matrix(rnorm(6, 10, 5), ncol = 3)
theta <- matrix(c(-2, 1, -2, 1), ncol = 2)
x <- rmixsespc(200, probs, mu, theta)$x
b <- dmixsespc(x, probs, mu, theta)
## the lat and long are decreased to 10. Increase them back to 50 to
## see the difference
mixsespc.contour(probs, mu, theta, lat = 10, long = 10)
```

`Density of a mixture of elliptically symmetric distributions`
Density of a mixture of elliptically symmetric distributions

Description

Density of a mixture of elliptically symmetric distributions.

Usage

```
dmixsespc(y, probs, mu, theta, logden = FALSE)
dmixesag(y, probs, mu, gam, logden = FALSE)
```

Arguments

`y`	A matrix with unit vectors.
`probs`	This is avector with the mixing probability of each group.
`mu`	A matrix with the mean direction of each group.
`theta`	A matrix with the θ parameters of the SESPC distribution of each group.
`gam`	A matrix with the γ parameters of the ESAG distribution of each group.
`logden`	If you the logarithm of the density values set this to TRUE.

Details

The function computes the density for a given mixture of SESPC or ESAG distributions.

Value

A vector with the (log) density values of y.

Author(s)

Michail Tsagris.

R implementation and documentation: Michail Tsagris <mtsagris@uoc.gr>.

References

Perdikis T., Alharbi N. and Tsagris M. (2026). Model–based clustering for spherical and hyper–spherical data using elliptically symmetric distributions.

https://arxiv.org/abs/2605.27496

See Also

mixsespc.mle, rmixsespc

Examples

```
probs <- c(0.5, 0.5)
mu <- matrix(rnorm(6, 10, 5), ncol = 3)
theta <- matrix(c(-2, 1, -2, 1), ncol = 2)
x <- rmixsespc(200, probs, mu, theta)$x
b <- dmixsespc(x, probs, mu, theta)
```

---

geodesic_pca_sphere *Principal geodesic analysis based subsphere projection*

Description

Principal geodesic analysis based subsphere projection.

Usage

```
geodesic_pca_sphere(x, k = 2)
```

Arguments

x	A matrix containing directional data in the hyper-sphere.
k	The dimensionality to which the data will be projected. For the sphere this should 2 (default value).

Details

We use a geodesic-aware dimensionality reduction that combines Principal Geodesic Analysis (Fletcher et al. 2004) with projection to a lower-dimensional sphere, similar in spirit to Principal Nested Spheres (Jung et al. 2012) but using global optimization rather than iterative reduction.

Value

A matrix with the projected data onto the sphere (or a lower dimension hyper-sphere).

Author(s)

Michail Tsagris.

R implementation and documentation: Michail Tsagris <mtsagris@uoc.gr>.

References

Jung S., Dryden I. L. and Marron J. S. (2012). Analysis of principal nested spheres. Biometrika, 99(3): 551–568.

Fletcher P. T., Lu C., Pizer S. M. and Joshi S. (2004). Principal geodesic analysis for the study of nonlinear statistics of shape. IEEE Transactions on Medical Imaging, 23(8): 995–1005.

See Also

sphereplot

Examples

```
x <- as.matrix( iris[, 1:4] )
x <- x / sqrt( rowSums(x^2) )
y <- geodesic_pca_sphere(x)
sphereplot(y)
```

`Interactive 3D plot of spherical data`

Interactive 3D plot of spherical data

Description

Interactive 3D plot of spherical data.

Usage

```
sphereplot(dat, col = NULL, bgcol = "snow")
```

Arguments

`dat`	A matrix with three columns, unit-vectors, spherical data.
`col`	If you want the points to appear with different colours put numbers here, otherwise leave it NULL.
`bgcol`	The color of the surface of the sphere.

Value

An interactive 3D plot of the spherical data will appear.

Author(s)

Michail Tsagris.

R implementation and documentation: Michail Tsagris <mtsagris@uoc.gr>.

See Also

mixsespc.contour

Examples

```
x <- as.matrix( iris[, 1:3] )
x <- x / sqrt( rowSums(x^2) )
sphereplot(x, col = as.numeric(iris[, 5]))
```

Mixtures of elliptically symmetric distributions

Mixtures of elliptically symmetric distributions

Description

It performs model based clustering for spherical data assuming ellitpically symetric distributions.

Usage

```
mixsespc.mle(x, g = 2, tol = 1e-4)
mixesag.mle(x, g = 2, tol = 1e-4)
```

Arguments

x	A matrix with the data expressed as unit vectors.
g	The number of groups to fit. It must be greater than or equal to 2.
tol	The tolerance value to terminate the EM algorithm.

Details

The functions performs model-based clustering using mixtures of SESPC or mixtures of ESAG distributions. The initial step of the algorithm is based on a Gaussian mixture model.

Value

A list including:

`param`	A matrix with the mean direction, the concentration parameters and mixing probability of each group.
`loglik`	The value of the maximised log-likelihood.
`probs`	The estimated probabilities of each observation to belong to each cluster.
`pred`	The predicted group of each observation.
`iter`	The number of iteration required by the EM algorithm.
`runtime`	The run time of the algorithm. A numeric vector. The first element is the user time, the second element is the system time and the third element is the elapsed time.

Author(s)

Michail Tsagris and Theodoros Perdikis.

R implementation and documentation: Michail Tsagris <mtsagris@uoc.gr>.

References

Perdikis T., Alharbi N. and Tsagris M. (2026). Model–based clustering for spherical and hyper–spherical data using elliptically symmetric distributions.

https://arxiv.org/abs/2605.27496

See Also

`bic.mixsespc`

Examples

```
x <- as.matrix( iris[, 1:3] )
x <- x / sqrt( rowSums(x^2) )
mixsespc.mle(x)
```

`Simulation of random values from a mixture of elliptically symmetric distributions`

Simulation of random values from a mixture of elliptically symmetric distributions

---

Description

The function simulates random values simulation from a given mixture of elliptically symmetric distributions.

Usage

```
rmixsespc(n, probs, mu, theta)
rmixesag(n, probs, mu, gam)
```

Arguments

n	The sample size.
probs	This is avector with the mixing probability of each group.
mu	A matrix with the mean direction of each group.
theta	A matrix with the θ parameters of the SESPC distribution of each group.
gam	A matrix with the γ parameters of the ESAG distribution of each group.

Details

The function simulates random values simulation from a given mixture of SESPC or a mixture of ESAG distributions.

Value

A list including:

id	An indicator of the group of each simulated vector.
x	A matrix with the simulated data.

Author(s)

Michail Tsagris and Theodoros Perdiki.

R implementation and documentation: Michail Tsagris <mtsagris@uoc.gr> and Theodoros Perdikis <theoperdikis92@gmail.com>.

References

Perdikis T., Alharbi N. and Tsagris M. (2026). Model–based clustering for spherical and hyper–spherical data using elliptically symmetric distributions.

https://arxiv.org/abs/2605.27496

See Also

mixsespc.mle, bic.mixsespc

Examples

```
probs <- c(0.5, 0.5)
mu <- matrix(rnorm(6, 10, 5), ncol = 3)
theta <- matrix( c(-2, 1, -2, 1), ncol = 2 )
x <- rmixsespc(200, probs, mu, theta)$x
```

Index