Package: navmix (via r-universe)

October 5, 2024

Type Package

Title A package for performing directional clustering by fitting a noise-augmented von Mises-Fisher mixture model

Version 0.2.1

Author Authors@R: person(``Andrew'', ``Grant'', email = ``andrew.grant@mrc-bsu.cam.ac.uk'', role = c(``aut'', ``cre''))

Maintainer The package maintainer <yourself@somewhere.net>

Description A package for performing directional clustering by fitting a noise-augmented von Mises-Fisher mixture model

License GPL-3

Encoding UTF-8

LazyData true

RoxygenNote 7.1.1

Suggests knitr, rmarkdown

VignetteBuilder knitr

URL https://github.com/aj-grant/navmix

Imports expm, ggplot2, plotrix, RColorBrewer, skmeans, tidyr

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

RemoteUrl https://github.com/aj-grant/navmix

RemoteRef HEAD

RemoteSha c3feac409f14f1bd622a4363dbe9403c558f3b79

Contents

---

navmix   *Noise-augmented directional clustering*

Description

Performs directional clustering by fitting a noise-augmented von Mises-Fisher mixture model

Usage

```
navmix(
  x,
  K = 10,
  select_K = TRUE,
  common_kappa = FALSE,
  pj_ini = 0.05,
  no_ini = 5,
  tol = 1e-04,
  max_iter = 100,
  plot = FALSE,
  plot_heat = TRUE,
  plot_heat_mu = FALSE,
  plot_parallel = TRUE,
  plot_radial = FALSE,
 plot_radial_options = list(plot_radial_separate = FALSE, radial_legend_pos = c(-2.5,
    2.7), radial_separate_col = 2)
)
```

Arguments

x	Matrix of values where rows represent observations and columns represent features.
K	The number of clusters to fit.
select_K	If TRUE (the default setting), the number of clusters will be chosen by BIC, with K the maximum number of clusters considered. If FALSE, then a model with K clusters will be fit.
pj_ini	The initial proportion of observations which belong in the noise cluster. Must be a number greater or equal to 0 and strictly less than 1. The default value is 0.05. If set to 0, no observations will be placed in the noise cluster.
no_ini	The number of time the algorithm is run with different initialisations. Must be a number greater than zero. The default value is 5.
tol	The tolerance threshold for convergence of the EM algorithm. Must be a number greater than 0. The default value is 1.0e-4.
max_iter	The maximum number of iterations of the EM algorithm. Must be a number greater than 0. The default value is 100.
plot	Plots of the results will be produced if set to TRUE. Default is FALSE.

plot_heat	Produces a heatmap of the results if plot is set to TRUE. The heatmap will also be returned as a ggplot object.
plot_radial	Produces (a) radial plot(s) of the results if plot is set to TRUE.
common_kapp	If TRUE, then model will force the kappa parameter to be equal for all clusters, except the noise cluster.

plot_radial_separate

If set to FALSE (the default value), the fitted means of each cluster are plotted on the same radial plot. If set to TRUE, they are plotted on separate radial plots.

radial_legend_pos

Adjusts the position of the legend for a radial plots with all fitted means plotted together.

radial_separate_col

Adjusts the format of the output of radial plots on separate plots.

Value

Returned are the BIC values for each model fitted ($BIC), the final fitted model ($fit) and, if produced, the heatmap as a ggplot object ($heatmap_plot). The fitted model has the following.

mu	A matrix where each column represents the mean of the fitted von Mises-Fisher distribution for each cluster.
kappa	A row vector where each element represents the kappa parameter of the fitted von Mises-Fisher distribution for each cluster.
g	A matrix of probabilities for each observation belonging to each cluster. The value in the jth row and kth column represents the probability that the jth observation belongs to the kth cluster.
z	A vector of the cluster membership of each observation when allocated according to the cluster for which it has the highest probability of membership (hard clustering).
bic	The BIC for the fitted model.
l	The value of the likelihood function at the estimated parameters.

Index