Package ‘shapeNA’

October 15, 2020

Title M-Estimation of Shape for Data With Missing Values

Version 0.0.0.9000

Description What the package does (one paragraph).

License `use_mit_license()`, `use_gpl3_license()` or friends to pick a license

Encoding UTF-8

LazyData true

Roxygen list(markdown = TRUE)

RoxygenNote 7.1.1

Suggests knitr, rmarkdown

VignetteBuilder knitr

RemoteType github

RemoteHost api.github.com

RemoteRepo shapeNA

RemoteUsername kathariemer

RemoteRef HEAD

RemoteSha dada4e4a304835c891655a4f2153395984c6de0f

GithubRepo shapeNA

GithubUsername kathariemer

GithubRef HEAD

GithubSHA1 dada4e4a304835c891655a4f2153395984c6de0f

NeedsCompilation no

Author Katharina Riemer [cre, aut],
Gabriel Frahm [aut] (<https://orcid.org/0000-0001-7507-730X>),
Klaus Nordhausen [aut] (<https://orcid.org/0000-0002-3758-8501>),
Una Radojicic [aut] (<https://orcid.org/0000-0003-0329-0595>)

Maintainer Katharina Riemer <maintainer@email.com>

R topics documented:

barplot.naBlocks 2
plot.naBlocks 2
plot.shapeNA 3
powerShape 3
powerShapeNA 4
print.naBlocks 6
print.shapeNA 6
print.summary.shapeNA 7
summary.shapeNA 7

Index **8**

`barplot.naBlocks` *Barplot showcasing missingness proportion of the original data*

Description

Barplot showcasing missingness proportion of the original data

Usage

```
## S3 method for class 'naBlocks'
barplot(obj, sortNA = FALSE)
```

`plot.naBlocks` *plot missingness pattern of data*

Description

plot missingness pattern of data

Usage

```
## S3 method for class 'naBlocks'
plot(x, orderProp = TRUE, ...)
```

`plot.shapeNA` *Crude visualization of shape estimate*

Description

If estimate comes from missing data, additionally the columns are marked with a colored bar, indicating their missingness proportion

Usage

```
## S3 method for class 'shapeNA'
plot(obj, legend = TRUE, message = TRUE)
```

`powerShape` *Computing M-Estimators of Shape for Data Without Missing Values*

Description

`powerShape`, `tylerShape` and `classicShape` compute power M-estimators of shape. Using `powerShape` and `classicShape` it is even possible to compute M-estimators of covariance matrices. These functions also compute estimates of location if no `center` is supplied.

Usage

```
powerShape(x, alpha, center = NULL, normalization = c("det", "trace", "one"), maxiter = 1e4, eps = 1e-6)

tylerShape(x, center = NULL, normalization = c("det", "trace", "one"), maxiter = 1e4, eps = 1e-6)

classicShape(x, center = NULL, normalization = c("det", "trace", "one"), maxiter = 1e4, eps = 1e-6)
```

Arguments

`x` A numeric data matrix or data.frame without missing data.

`alpha` Tail index, a numeric value from the interval 0, 1. Determines the power function. For more information see ’Details’.

`center` An optional vector of the data’s center, if NULL the center will be estimated simultaneously to the shape estimate.

`normalization` A string, determines scale of returned shape estimate. The possible values are

- ’det’ s.t. the returned shape estimate has determinant 1.
- ’trace’ s.t. the returned shape estimate has trace `p`.
- ’one’ s.t. the returned shape estimate’s first entry is 1.

`maxiter` A positive integer, restricting the maximum number of iterations.

`eps` A numeric, specifying tolerance level of when the iteration stops.

Details

For multivariate normally distributed data, `classicShape` is an ML-estimator. This is a special case of the power M-estimator with tail index `alpha` = 0 and returns the empirical covariance matrix and the empirical mean vector.

`tylerShape` maximizes the likelihood function after projecting the observed data of each individual onto the unit hypersphere, in which case we obtain an angular central Gaussian distribution. This is a special case of the power M-estimator with tail index `alpha` = 1 and returns Tyler's M-estimator of scatter and an affine equivariant multivariate median.

`powerShape` requires an additional parameter, the so-called tail index `alpha`. For asymptotic normality, this index should be chosen taking into consideration the data. For heavy tailed data, the index should be closer to 1, for light tailed data the index should be chosen closer to 0.

Value

A `shapeNA` object with a shape estimate `S` and either a center `mu`, which was either the supplied `center` vector or has been estimated.

a shapeNA object, which contains a shape and center estimate

References

Frahm, G., & Jaekel, U. (2010). A generalization of Tyler's M-estimators to the case of incomplete data. Computational Statistics & Data Analysis, 54(2), 374-393.

Frahm, G., Nordhausen, K., & Oja, H. (2020). M-estimation with incomplete and dependent multivariate data. Journal of Multivariate Analysis, 176, 104569.

See Also

powerShapeNA

tylerShapeNA

classicShapeNA

Examples

```
x <- mvtnorm::rmvt(100, toeplitz(seq(1, 0.1, length.out=5)))
res <- powerShape(x, alpha=0.67, normalization='one')
```

`powerShapeNA` *Computing M-Estimators of Shape for Data With Missing Values*

Description

`powerShapeNA`, `tylerShapeNA` and `classicShapeNA` compute power M-estimators of shape for data with missing values. The underlying missingness mechanism should be MCAR. These functions also compute estimates of location if no `center` is supplied.

Usage

```
powerShapeNA(x, alpha, center = NULL, normalization = c("det", "trace", "one"),
         maxiter = 1e4, eps = 1e-6)

tylerShapeNA(x, center = NULL, normalization = c("det", "trace", "one"),
          maxiter = 1e4, eps = 1e-6)

classicShapeNA(x, center = NULL, normalization = c("det", "trace", "one"),
         maxiter = 1e4, eps = 1e-6)
```

Arguments

x — A data matrix or data.frame with missing data and p > 2 columns. Representing sample from generalized elliptical distribution and MCAR missingness

alpha — Tail index, a numeric value from the interval 0, 1. Determines the power function. For more information see 'Details'.

center — An optional vector of the data's center, if NULL the center will be estimated simultaneously to the shape estimate.

normalization — A string, determines scale of returned shape estimate. The possible values are

- 'det' s.t. the returned shape estimate has determinant 1.
- 'trace' s.t. the returned shape estimate has trace p.
- 'one' s.t. the returned shape estimate's first entry is 1.

maxiter — A positive integer, restricting the maximum number of iterations.

eps — A numeric, specifying tolerance level of when the iteration stops.

Details

For multivariate normally distributed data, classicShapeNA is an ML-estimator. This is a special case of the power M-estimator with tail index alpha = 0 and returns the empirical covariance matrix and the empirical mean vector.

tylerShapeNA maximizes the likelihood function after projecting the observed data of each individual onto the unit hypersphere, in which case we obtain an angular central Gaussian distribution. This is a special case of the power M-estimator with tail index alpha = 1 and returns Tyler's M-estimator of scatter and an affine equivariant multivariate median.

powerShapeNA requires an additional parameter, the so-called tail index alpha. For asymptotic normality, this index should be chosen taking into consideration the data. For heavy tailed data, the index should be closer to 1, for light tailed data the index should be chosen closer to 0.

Value

A shapeNA object with a shape estimate S and either a center mu, which was either the supplied center vector or has been estimated.

References

Frahm, G., & Jaekel, U. (2010). A generalization of Tyler's M-estimators to the case of incomplete data. Computational Statistics & Data Analysis, 54, 374-393. doi:10.1016/j.csda.2009.08.019.

Frahm, G., Nordhausen, K., & Oja, H. (2020). M-estimation with incomplete and dependent multivariate data. Journal of Multivariate Analysis, 176, 104569. doi:10.1016/j.jmva.2019.104569.

Examples

```
## generate data set with missing values
x <- mvtnorm::rmvt(100, toeplitz(seq(1, 0.1, length.out = 3)), df = 5)
y <- mice::ampute(x, mech='MCAR')$amp
## compute M-estimate
res <- powerShapeNA(y, alpha = 0.5)
summary(res)
```

print.naBlocks *Print missingness pattern*

Description

Print missingness pattern

Usage

```
## S3 method for class 'naBlocks'
print(obj)
```

print.shapeNA *print method for elements of class shapeNA*

Description

Only print M-estimates and alpha level

Usage

```
## S3 method for class 'shapeNA'
print(obj)
```

`print.summary.shapeNA` *print method for class summary.shapeNA*

Description

print method for class summary.shapeNA

Usage

```
## S3 method for class 'summary.shapeNA'
print(obj, ...)
```

Arguments

`obj`	object returned from summary.shapeNA
`...`	further arguments

Value

invisibly return NULL

`summary.shapeNA` *summary method for class shapeNA*

Description

summary method for class shapeNA

Usage

```
## S3 method for class 'shapeNA'
summary(obj, ...)
```

Arguments

`obj`	an object of class shapeNA, usually from a call to powerShape or similar functions
`...`	further arguments

Value

object of class shapeNA

Examples

```
obj <- tylerShape(mvtnorm::rmvt(100, diag(3)))
summary(obj)
```

Index

0, 1, *3*, *5*

barplot.naBlocks, 2

classicShape *(*powerShape*)*, 3
classicShapeNA *(*powerShapeNA*)*, 4

plot.naBlocks, 2
plot.shapeNA, 3
powerShape, 3
powerShapeNA, 4
print.naBlocks, 6
print.shapeNA, 6
print.summary.shapeNA, 7

summary.shapeNA, 7

tylerShape *(*powerShape*)*, 3
tylerShapeNA *(*powerShapeNA*)*, 4