Package 'REFA'

November 19, 2023

Type Package

Title Robust Exponential Factor Analysis

Date 2023-11-01

Version 0.1.0

Author Jiaqi Hu [cre, aut],
Xueqin Wang [aut]

Maintainer Jiaqi Hu <hujiaqi@mail.ustc.edu.cn>

Description

A robust alternative to the traditional principal component estimator is proposed within the framework of factor models, known as Robust Exponential Factor Analysis, specifically designed for the modeling of high-dimensional datasets with heavy-tailed distributions. The algorithm estimates the latent factors and the loading by minimizing the exponential squared loss function. To determine the appropriate number of factors, we propose a modified rank minimization technique, which has been shown to significantly enhance finite-sample performance.

Imports mvtnorm

Depends R (>= 3.5.0)

License GPL-3

Encoding UTF-8

RoxygenNote 7.2.3

NeedsCompilation no

Repository CRAN

Date/Publication 2023-11-19 15:20:05 UTC

R topics documented:

ECC *Estimation of errors for common component*

Description

Estimation of errors for common component

Usage

```
ECC(Chat, C)
```

Arguments

`Chat`	The estimated common component
`C`	The true common component

Value

a numeric value of the ECC

Author(s)

Jiaqi Hu

References

Manuscript: Robust factor analysis with exponential squared loss

Examples

```
dat = gendata()
Y = dat$Y
F0 = dat$F0
L0 = dat$L0
C0 = F0
res = REFA(dat$Y, r = 3)
Fhat = res$Fhat
Lhat = res$Lhat
Chat = Fhat
ECC(Chat, C0)
```

est_num *Estimating Factor Numbers Corresponding PCA*

Description

Estimating Factor Numbers Corresponding PCA

Usage

```
est_num(X, kmax = 8, type = "BIC3")
```

Arguments

X	Input matrix, of dimension $T \times N$. Each row is an observation with N features at time point t.
kmax	The user-supplied maximum factor numbers.
type	the method used.

Value

the estimated factor numbers

Author(s)

Jiaqi Hu

References

Manuscript: Robust factor analysis with exponential squared loss

Examples

```
dat = gendata()
est_num(dat$Y)
```

FA *Principal Component Analysis for Factor Models*

Description

Principal Component Analysis for Factor Models

Usage

```
FA(X, r)
```

Arguments

X	Input matrix, of dimension $T \times N$. Each row is an observation with N features at time point t.
r	A positive integer indicating the factor numbers.

Value

Fhat	The estimated factor matrix.
Lhat	The estimated loading matrix.

Author(s)

Jiaqi Hu

References

Manuscript: Robust factor analysis with exponential squared loss

Examples

```
##---- Should be DIRECTLY executable !! ----
```

gendata *Data generation process*

Description

Generate heavy-tailed data.

Usage

```
gendata(seed = 1, T = 50, N = 50, type = "1a")
```

Arguments

`seed` the seed used in the data generation process.

`T` time dimension.

`N` cross-sectional dimension.

`type` the type of the data generation process, it can be `"1a"`, `"1b"`, `"1c"`, `"1d"`, `"2a"`, `"2b"`, `"2c"`, `"2d"`.

Value

a list consisting of `Y`, `F0`, `L0`.

Author(s)

Jiaqi Hu

References

Manuscript: Robust factor analysis with exponential squared loss

Examples

```
dat = gendata()
Y = dat$Y
head(Y)
```

REFA *Robust Exponential Factor Analysis*

Description

Robust Exponential Factor Analysis

Usage

```
REFA(Y, r = 3, tau = 0.75, q = 0.05, eps = 1e-05, init = TRUE)
```

Arguments

`Y`	Input matrix, of dimension $T \times N$. Each row is an observation with N features at time point t.
`r`	A positive integer indicating the factor numbers.
`q`	Hyper parameter
`eps`	The stopping criterion parameter. The default is 1e-5.
`tau`	Hyper parameter
`init`	Warn start of the algorithm. If `init = TRUE`, use modified PCA initialization. If `init` is a list contains `F0` and `L0`, we will use this initialization. Otherwise, use traditional PCA initialization.

Value

`Fhat`	The estimated factor matrix.
`Lhat`	The estimated loading matrix.
`loss`	the value of the loss function.

Author(s)

Jiaqi Hu

References

Manuscript: Robust factor analysis with exponential squared loss

Examples

```
dat = gendata()
REFA(dat$Y, r = 3)
```

REFA_FN *Estimating Factor Numbers via Modified Rank Minimization*

Description

Estimating Factor Numbers via Modified Rank Minimization

Usage

```
REFA_FN(Y, rmax = 8, tau = 0.75, q = 0.1, eps = 1e-04, init = TRUE)
```

Arguments

`Y` Input matrix, of dimension $T \times N$. Each row is an observation with N features at time point t.

`rmax` The bound of the number of factors.

`q` Hyper parameter in modified PCA algorithm. Default is `0.05`.

`eps` The stopping criterion parameter. Default is `1e-5`.

`tau` Hyper parameter in selecting γ of the loss function.

`init` Warn start by modified PCA algorithm. Default is `TRUE`.

Value

`rhat` The estimated factor number.

`Fhat` The estimated factor matrix.

`Lhat` The estimated loading matrix.

`loss` the value of the loss function.

Author(s)

Jiaqi Hu

References

Manuscript: Robust factor analysis with exponential squared loss

Examples

```
dat = gendata()
REFA_FN(dat$Y, rmax = 8)
```

TR *Trace ratios*

Description

Trace ratios

Usage

```
TR(Fhat, F0)
```

Arguments

Fhat	The estimated factors.
F0	The true factors.

Value

a numeric value of the trace ratios.

Author(s)

Jiaqi Hu

References

Manuscript: Robust factor analysis with exponential squared loss

Examples

```
dat = gendata()
Y = dat$Y
F0 = dat$F0
res = REFA(dat$Y, r = 3)
Fhat = res$Fhat
TR(Fhat, F0)
```

Index