

Package: RMFM (via r-universe)

November 27, 2024

Type Package

Title Robust Matrix Factor Model

Version 1.1.0

Date 2024-09-26

Author Wei Liu [aut, cre], Xiaolu Jiang [aut], Jinyu Nie [aut]

Maintainer Wei Liu <liuwei8@scu.edu.cn>

Description We introduce a robust matrix factor model that explicitly incorporates tail behavior and employs a mean-shift term to avoid efficiency losses through pre-centering of observed matrices. More details on the methods related to our paper are currently under submission. A full reference to the paper will be provided in future versions once the paper is published.

License GPL-3

LazyData true

Depends irlba, R (>= 3.5.0)

Imports stats, LaplacesDemon, MixMatrix, COAP, Rcpp (>= 1.0.10)

Suggests knitr, rmarkdown

LinkingTo Rcpp, RcppArmadillo

Encoding UTF-8

RoxygenNote 7.3.2

Repository <https://feiyong.r-universe.dev>

RemoteUrl <https://github.com/feiyong/rmfm>

RemoteRef HEAD

RemoteSha f293dfb4a1766f3782bec67c0724a0e2c2e41113

Contents

ER.RMFM	2
gendata_rmfm	3
RMFM	4

Index	7
--------------	----------

ER.RMFM

*Select the structure dimension of factor matrix***Description**

Select the structure dimension of factor matrix in the high-dimensional robust matrix factor model

Usage

```
ER.RMFM(
  X,
  r_max = 10,
  epsELBO = 1e-09,
  maxIter = 20,
  verbose = FALSE,
  seed = 1
)
```

Arguments

X	a $p_1 \times p_2 \times T$ array, which is the observed matrix from each individual, where T is the sample size.
r_max	an optional positive integer, specify the upper bound of row and column factors; default as 10.
epsELBO	an optional positive value, tolerance of relative variation rate of the variational lower bound value, default as '1e-9'.
maxIter	the maximum iteration of the VEM algorithm. The default is 30.
verbose	a logical value, whether output the information in iteration.
seed	an optional integer, specify the random seed for reproducibility in initialization.

Details

None

Value

return a list including the following components:

- rvec - a two-dimensional vector, the estimated row and column numbers of factors.
- svrMat - a r_max-by-2 matrix, the singular value ratios.

References

None

See Also

None

Examples

```

r1 <- 4; r2 <- 3;
Tt <- 100; type <- 'MatrixT'
p1 <- 50; p2 <- 50
datlist <- gendata_rmfm(i = 1, Tt = Tt, p1 = p1, p2 = p2, r1 = r1, r2 = r2,
                      rho = 1, type = 'MatrixT', nu = 3)
str(datlist)
res <- ER.RMFM(datlist$X, r_max = 10, epsELBO = 1e-9, maxIter = 10, verbose = FALSE, seed = 1)
res

```

gendata_rmfm

Generate simulated data

Description

Generate simulated data from robust matrix factor models

Usage

```

gendata_rmfm(
  i = 1,
  Tt = 100,
  p1 = 50,
  p2 = 40,
  r1 = 4,
  r2 = 3,
  rho = 0.01,
  type = c("MatrixT", "MatrixN"),
  nu = 1
)

```

Arguments

<code>i</code>	a positive integer, the random seed for reproducibility of data generation process.
<code>Tt</code>	a positive integer, specify the sample size.
<code>p1</code>	a positive integer, specify the row dimension of the observed matrix.
<code>p2</code>	a positive integer, specify the column dimension of the observed matrix.
<code>r1</code>	a positive integer, specify the number of row factors; default as 4
<code>r2</code>	a positive integer, specify the number of column factors; default as 3.
<code>rho</code>	a positive real, specify the signal strength of factor matrices.
<code>type</code>	a string, specify the type of error matrix, default as <code>type = 'MatrixN'</code> ; supportint matrix t distribution 'MatrixT' and matrix normal distribution 'MatrixN'.
<code>nu</code>	a positive integer, specify the degree freedom of the matrix t distribution when <code>type = 'MatrixT'</code> .

Value

return a list including the following components:

- X - $p1 * p2 * T$ array, which is the observed matrix from each individual, where T is the sample size.
- CC - $p1 * p2 * T$ array, which is the common component matrix for each individual.
- $F0$ - $r1 * r2 * T$ array, which is the generated factor matrix for each individual, where T is the sample size.
- $R0$ - a $p1$ -by- $r1$ matrix, the row loading matrix.
- $C0$ - a $p2$ -by- $r2$ matrix, the column loading matrix.
- μ_0 - a $p1$ -by- $p2$ matrix, the mean matrix.

Examples

```
r1 <- 4; r2 <- 3;
Tt <- 100; type <- 'MatrixT'
p1 <- 100; p2 <- 50
datlist <- gendata_rmfm(i = 1, Tt = Tt, p1 = p1, p2 = p2, r1 = r1, r2 = r2,
                      rho = 0.01, type = type, nu = 1)
str(datlist)
```

 RMFM

Fit the high-dimensional robust matrix factor model

Description

Fit the high-dimensional robust matrix factor model via variational inference.

Usage

```
RMFM(
  X,
  r1 = 10,
  r2 = 10,
  epsELBO = 1e-09,
  maxIter = 30,
  verbose = TRUE,
  seed = 1,
  cal_eigs = FALSE
)
```

Arguments

X	a $p1 * p2 * T$ array, which is the observed matrix from each individual, where T is the sample size.
r1	an optional positive integer, specify the number of row factors; default as 10.
r2	an optional positive integer, specify the number of column factors; default as 10.
epsELBO	an optional positive value, tolerance of relative variation rate of the variational lower bound value, default as '1e-9'.
maxIter	the maximum iteration of the VEM algorithm. The default is 30.
verbose	a logical value, whether output the information in iteration.
seed	an optional integer, specify the random seed for reproducibility in initialization.
cal_eigs	an optional logical value, specify whether calculate the eigenvalues of covariance matrix, default as FALSE.

Details

None

Value

return a list including the following components:

- hF - a $r1 * r2 * T$ array, which is the estimated factor matrix for each individual, where T is the sample size.
- hmu - a $p1$ -by- $p2$ matrix, the estimated mean matrix.
- hR - the estimated row loading matrix.
- hC - the estimated column loading matrix.
- hnu - the estimated degree of freedom for the error term.
- hLambda1 - a $p1$ vector, the estimated row scatter matrix for error.
- hLambda2 - a $p2$ vector, the estimated column scatter matrix for error.
- dR - NULL if cal_eigs=FALSE; a group of eigenvalues of the sample covariance across rows if cal_eigs=TRUE.
- dC - NULL if cal_eigs=FALSE; a group of eigenvalues of the sample covariance across columns if cal_eigs=TRUE.
- ELBO - the ELBO value when algorithm stops;
- ELBO_seq - the sequence of ELBO values.
- time_use - the running time in model fitting of RMFM;

References

None

See Also

None

Examples

```
r1 <- 4; r2 <- 3;
Tt <- 100; type <- 'MatrixT'
p1 <- 50; p2 <- 50
datlist <- gendata_rmfm(i = 1, Tt = Tt, p1 = p1, p2 = p2, r1 = r1, r2 = r2,
                      rho = 1, type = 'MatrixT', nu = 1)
str(datlist)
reslist <- RMFM(X = datlist$X, r1 = r1, r2 = r2, verbose = TRUE, maxIter = 6)
```

Index

ER.RMFM, [2](#)

gendata_rmfm, [3](#)

RMFM, [4](#)