Package ‘sEparaTe’

June 27, 2021

Title Maximum Likelihood Estimation and Likelihood Ratio Test Functions for Separable Variance-Covariance Structures

Version 0.3.0

Maintainer Timothy Schwinghamer <timothy.schwinghamer@AGR.GC.CA>

Description Maximum likelihood estimation of the parameters of matrix and 3rd-order tensor normal distributions with unstructured factor variance covariance matrices, two procedures, and for unbiased modified likelihood ratio testing of simple and double separability for variance-covariance structures, two procedures.

Depends R (>= 3.1.0)

License MIT + file LICENSE

Encoding UTF-8

LazyData true

RoxygenNote 7.1.1

Suggests knitr, rmarkdown

NeedsCompilation no

Author Ameur Manceur [aut],
Timothy Schwinghamer [aut, cre],
Pierre Dutilleul [aut, cph]

Repository CRAN

Date/Publication 2021-06-27 07:40:02 UTC

R topics documented:

data2d — *Two dimensional data set*

Description

An i.i.d. random sample of size 7 from a 2 x 3 matrix normal distribution, for a small numerical example of the use of the functions mle2d_svc and lrt2d_svc from the **sEparaTe** package

Usage

```
data2d
```

Format

A frame (excluding the headings) with 42 lines of data and 4 variables:

K an integer ranging from 1 to 7, the size of an i.i.d. random sample from a 2 x 3 matrix normal distribution

Id1 an integer ranging from 1 to 2 , the number of rows of the matrix normal distribution

Id2 an integer ranging from 1 to 3 , the number of columns of the matrix normal distribution

value2d the sample data for the observed variable

data3d — *Three dimensional data set*

Description

An i.i.d. random sample of size 11 from a 2 x 2 x 2 tensor normal distribution, for a small numerical example of the use of the functions mle3d_svc and lrt3d_svc from the **sEparaTe** package

Usage

```
data3d
```

Format

A frame (excluding the headings) with 88 lines of data and 5 variables:

K an integer ranging from 1 to 11, the size of an i.i.d. random sample from a 2 x 2 x 2 tensor matrix normal distribution

Id3 an integer ranging from 1 to 2, the number of rows of the 3rd-order tensor normal distribution

Id4 an integer ranging from 1 to 2, the number of columns of the 3rd-order tensor normal distribution

Id5 an integer ranging from 1 to 2, the number of edges of the 3rd-order tensor normal distribution

value3d the sample data for the observed variable `lrt2d_svc` *Unbiased modified likelihood ratio test for simple separability of a variance-covariance matrix.*

Description

A likelihood ratio test (LRT) for simple separability of a variance-covariance matrix, modified to be unbiased in finite samples. The modification is a penalty-based homothetic transformation of the LRT statistic. The penalty value is optimized for a given mean model, which is left unstructured here. In the required function, the Id1 and Id2 variables correspond to the row and column subscripts, and are the second and third columns in the matrix (2d) data file, respectively; "value2d" refers to the observed variable, and is the fourth column in the matrix data file.

Usage

```
lrt2d_svc(
  formula_2d,
  subject,
  data_2d = list(),
  eps,
  maxiter,
  startmat,
  sign.level,
  n.simul
)
```

Arguments

`formula_2d`	value2d~Id1+Id2
`subject`	the replicate, also called the subject or individual, the first column in the matrix (2d) data file
`data_2d`	the name of the matrix data
`eps`	the threshold in the stopping criterion for the iterative mle algorithm (estimation)
`maxiter`	the maximum number of iterations for the mle algorithm (estimation)
`startmat`	the value of the second factor variance-covariance matrix used for initialization, i.e., to start the mle algorithm (estimation) and obtain the initial estimate of the first factor variance-covariance matrix
`sign.level`	the significance level, or rejection rate in the testing of the null hypothesis of simple separability for a variance-covariance structure, when the unbiased modified LRT is used, i.e., the critical value in the chi-square test is derived by simulations from the sampling distribution of the LRT statistic
`n.simul`	the number of simulations used to build the sampling distribution of the LRT statistic under the null hypothesis, using the same characteristics as the i.i.d. random sample from a matrix normal distribution

Output

"Convergence", TRUE or FALSE

"chi.df", the theoretical number of degrees of freedom of the asymptotic chi-square distribution that would apply to the unmodified LRT statistic for simple separability of a variance-covariance structure

"Lambda", the observed value of the unmodified LRT statistic

"critical.value", the critical value at the specified significance level for the chi-square distribution with "chi.df" degrees of freedom

"Decision.lambda" will indicate whether or not the null hypothesis of separability was rejected, based on the theoretical LRT statistic

"Simulation.critical.value", the critical value at the specified significance level that is derived from the sampling distribution of the unbiased modified LRT statistic

"Decision.lambda.simulation", the decision (acceptance/rejection) regarding the null hypothesis of simple separability, made using the theoretical (biased unmodified) LRT

"Penalty", the optimized penalty value used in the homothetic transformation between the biased unmodified and unbiased modified LRT statistics

"U1hat", the estimated variance-covariance matrix for the rows

"Standardized_U1hat", the standardized estimated variance-covariance matrix for the rows; the standardization is performed by dividing each entry of U1hat by entry(1, 1) of U1hat

"U2hat", the estimated variance-covariance matrix for the columns

"Standardized_U2hat", the standardized estimated variance-covariance matrix for the columns; the standardization is performed by multiplying each entry of U2hat by entry(1, 1) of U1hat

"Shat", the sample variance-covariance matrix computed from the vectorized data matrices

References

Manceur AM, Dutilleul P. 2013. Unbiased modified likelihood ratio tests for simple and double separability of a variance-covariance structure. Statistics and Probability Letters 83: 631-636.

Examples

```
output <- lrt2d_svc(value2d~Id1+Id2, subject = "K", data_2d = data2d, n.simul = 100)
output
```

`lrt3d_svc` *An unbiased modified likelihood ratio test for double separability of a variance-covariance structure.*

Description

A likelihood ratio test (LRT) for double separability of a variance-covariance structure, modified to be unbiased in finite samples. The modification is a penalty-based homothetic transformation of the LRT statistic. The penalty value is optimized for a given mean model, which is left unstructured here. In the required function, the Id3, Id4 and Id5 variables correspond to the row, column and edge subscripts, and are the second, third and fourth columns in the tensor (3d) data file, respectively; "value3d" refers to the observed variable, and is the fifth column in the tensor data file.

Usage

```
lrt3d_svc(
  formula_3d,
  subject,
  data_3d = list(),
  eps,
  maxiter,
  startmatU2,
  startmatU3,
  sign.level,
  n.simul
)
```

Arguments

`formula_3d`	value3d~Id3+Id4+Id5
`subject`	the replicate, also called individual, the first column in the tensor (3d) data file
`data_3d`	the name of the tensor data
`eps`	the threshold in the stopping criterion for the iterative mle algorithm (estimation)
`maxiter`	the maximum number of iterations for the mle algorithm (estimation)
`startmatU2`	the value of the second factor variance-covariance matrix used for initialization
`startmatU3`	the value of the third factor variance-covariance matrix used for initialization, i.e., startmatU3 together with startmatU2 are used to start the mle algorithm (estimation) and obtain the initial estimate of the first factor variance-covariance matrix U1
`sign.level`	the significance level, or rejection rate in the testing of the null hypothesis of simple separability for a variance-covariance structure, when the unbiased modified LRT is used, i.e., the critical value in the chi-square test is derived by simulations from the sampling distribution of the LRT statistic
`n.simul`	the number of simulations used to build the sampling distribution of the LRT statistic under the null hypothesis, using the same characteristics as the i.i.d. random sample from a tensor normal distribution

Output

"Convergence", TRUE or FALSE

"chi.df", the theoretical number of degrees of freedom of the asymptotic chi-square distribution that would apply to the unmodified LRT statistic for double separability of a variance-covariance structure "Lambda", the observed value of the unmodified LRT statistic

"critical.value", the critical value at the specified significance level for the chi-square distribution with "chi.df" degrees of freedom

"Decision.lambda", the decision (acceptance/rejection) regarding the null hypothesis of double separability, made using the theoretical (biased unmodified) LRT

"Simulation.critical.value", the critical value at the specified significance level that is derived from the sampling distribution of the unbiased modified LRT statistic

"Decision.lambda.simulation", the decision (acceptance/rejection) regarding the null hypothesis of double separability, made using the unbiased modified LRT

"Penalty", the optimized penalty value used in the homothetic transformation between the biased unmodified and unbiased modified LRT statistics

"U1hat", the estimated variance-covariance matrix for the rows

"Standardized_U1hat", the standardized estimated variance-covariance matrix for the rows; the standardization is performed by dividing each entry of U1hat by entry(1, 1) of U1hat

"U2hat", the estimated variance-covariance matrix for the columns

"Standardized_U2hat", the standardized estimated variance-covariance matrix for the columns; the standardization is performed by multiplying each entry of U2hat by entry(1, 1) of U1hat

"U3hat", the estimated variance-covariance matrix for the edges

"Shat", the sample variance-covariance matrix computed from the vectorized data tensors

References

Manceur AM, Dutilleul P. 2013. Unbiased modified likelihood ratio tests for simple and double separability of a variance-covariance structure. Statistics and Probability Letters 83: 631-636.

Examples

```
output <- lrt3d_svc(value3d~Id3+Id4+Id5, subject = "K", data_3d = data3d, n.simul = 100)
output
```

`mle2d_svc` *Maximum likelihood estimation of the parameters of a matrix normal distribution*

Description

Maximum likelihood estimation for the parameters of a matrix normal distribution **X**, which is characterized by a simply separable variance-covariance structure. In the general case, which is the case considered here, two unstructured factor variance-covariance matrices determine the covariability of random matrix entries, depending on the row (one factor matrix) and the column (the other factor matrix) where two **X**-entries are. In the required function, the Id1 and Id2 variables correspond to the row and column subscripts, and are the second and third columns in the matrix (2d) data file, respectively; “value” indicates the observed variable, and is the fourth column in the matrix data file.

Usage

```
mle2d_svc(formula_2d, subject, data_2d = list(), eps, maxiter, startmat)
```

Arguments

`formula_2d`	value2d~Id1+Id2
`subject`	the replicate, also called individual, the first column in the matrix (2d) data file
`data_2d`	the name of the matrix data
`eps`	the threshold in the stopping criterion for the iterative mle algorithm
`maxiter`	the maximum number of iterations for the iterative mle algorithm
`startmat`	the value of the second factor variance-covariance matrix used for initialization, i.e., to start the algorithm and obtain the initial estimate of the first factor variance-covariance matrix

Output

“Convergence”, TRUE or FALSE

“Iter”, will indicate the number of iterations needed for the mle algorithm to converge

“Xmeanhat”, the estimated mean matrix (i.e., the sample mean)

“First”, the row subscript, or the second column in the data file

“U1hat”, the estimated variance-covariance matrix for the rows

“Standardized.U1hat”, the standardized estimated variance-covariance matrix for the rows; the standardization is performed by dividing each entry of U1hat by entry(1, 1) of U1hat

“Second”, the column subscript, or the third column in the data file

“U2hat”, the estimated variance-covariance matrix for the columns

“Standardized.U2hat”, the standardized estimated variance-covariance matrix for the columns; the standardization is performed by multiplying each entry of U2hat by entry(1, 1) of U1hat

“Shat”, is the sample variance-covariance matrix computed from of the vectorized data matrices

References

Dutilleul P. 1990. Apport en analyse spectrale d’un periodogramme modifie et modelisation des series chronologiques avec repetitions en vue de leur comparaison en frequence. D.Sc. Dissertation, Universite catholique de Louvain, Departement de mathematique.

Dutilleul P. 1999. The mle algorithm for the matrix normal distribution. Journal of Statistical Computation and Simulation 64: 105-123.

Examples

```
output <- mle2d_svc(value2d~Id1+Id2, subject = "K", data_2d = data2d)
output
```

`mle3d_svc` *Maximum likelihood estimation of the parameters of a 3rd-order tensor normal distribution*

Description

Maximum likelihood estimation for the parameters of a 3rd-order tensor normal distribution **X**, which is characterized by a doubly separable variance-covariance structure. In the general case, which is the case considered here, three unstructured factor variance-covariance matrices determine the covariability of random tensor entries, depending on the row (one factor matrix), the column (another factor matrix) and the edge (remaining factor matrix) where two **X**-entries are. In the required function, the Id3, Id4 and Id5 variables correspond to the row, column and edge subscripts, and are the second, third and fourth columns in the tensor (3d) data file, respectively; "value3d" indicates the observed variable, and is the fifth column in the tensor data file.

Usage

```
mle3d_svc(
  formula_3d,
  subject,
  data_3d = list(),
  eps,
  maxiter,
  startmatU2,
  startmatU3
)
```

Arguments

`formula_3d`	value3d~Id3+Id4+Id5
`subject`	the replicate, also called individual, the first column in the tensor (3d) data file
`data_3d`	the name of the tensor data
`eps`	the threshold in the stopping criterion for the iterative mle algorithm
`maxiter`	the maximum number of iterations for the iterative mle algorithm
`startmatU2`	the value of the second factor variance covariance matrix used for initialization
`startmatU3`	the value of the third factor variance covariance matrix used for initialization, i.e., startmatU3 together with startmatU2 are used to start the algorithm and obtain the initial estimate of the first factor variance covariance matrix U1

Output

"Convergence", TRUE or FALSE

"Iter", the number of iterations needed for the mle algorithm to converge

"Xmeanhat", the estimated mean tensor (i.e., the sample mean)

"First", the row subscript, or the second column in the data file

"U1hat", the estimated variance-covariance matrix for the rows

"Standardized.U1hat", the standardized estimated variance-covariance matrix for the rows; the standardization is performed by dividing each entry of U1hat by entry(1, 1) of U1hat

"Second", the column subscript, or the third column in the data file

"U2hat", the estimated variance-covariance matrix for the columns

"Standardized.U2hat", the standardized estimated variance-covariance matrix for the columns; the standardization is performed by multiplying each entry of U2hat by entry(1, 1) of U1hat

"Third", the edge subscript, or the fourth column in the data file

"U3hat", the estimated variance-covariance matrix for the edges

"Shat", the sample variance-covariance matrix computed from the vectorized data tensors

Reference

Manceur AM, Dutilleul P. 2013. Maximum likelihood estimation for the tensor normal distribution: Algorithm, minimum sample size, and empirical bias and dispersion. Journal of Computational and Applied Mathematics 239: 37-49.

Examples

```
output <- mle3d_svc(value3d~Id3+Id4+Id5, subject = "K", data = data3d)
output
```

`sEparaTe` *MLE and LRT functions for separable variance-covariance structures*

Description

A package for maximum likelihood estimation (MLE) of the parameters of matrix and 3rd-order tensor normal distributions with unstructured factor variance-covariance matrices (two procedures), and for unbiased modified likelihood ratio testing (LRT) of simple and double separability for variance-covariance structures (two procedures).

Functions

mle2d_svc, for maximum likelihood estimation of the parameters of a matrix normal distribution

mle3d_svc, for maximum likelihood estimation of the parameters of a 3rd-order tensor normal distribution

lrt2d_svc, for the unbiased modified likelihood ratio test of simple separability for a variance-covariance structure

lrt3d_svc, for the unbiased modified likelihood ratio test of double separability for a variance-covariance structure

Data

data2d, a two-dimensional data set

data3d, a three-dimensional data set

References

Dutilleul P. 1999. The mle algorithm for the matrix normal distribution. Journal of Statistical Computation and Simulation 64: 105-123.

Manceur AM, Dutilleul P. 2013. Maximum likelihood estimation for the tensor normal distribution: Algorithm, minimum sample size, and empirical bias and dispersion. Journal of Computational and Applied Mathematics 239: 37-49.

Manceur AM, Dutilleul P. 2013. Unbiased modified likelihood ratio tests for simple and double separability of a variance covariance structure. Statistics and Probability Letters 83: 631-636.

Index