

Package: Rnest (via r-universe)

September 4, 2024

Type Package

Title Next Eigenvalue Sufficiency Test

Version 0.0.0.2

Description Determine the number of dimensions to retain in exploratory factor analysis. The main function, `nest()`, returns the solution and the `plot(nest())` returns a plot.

License MIT + file LICENSE

Encoding UTF-8

LazyData true

RoxygenNote 7.2.3

Imports crayon (>= 1.4.0), ggplot2 (>= 3.3.0), scales (>= 1.0.0), EFA.MRFA (>= 1.1.2), fungible (>= 2.3), MASS (>= 7.3-58.1)

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

RemoteUrl <https://github.com/quantmeth/rnest>

RemoteRef HEAD

RemoteSha 89e7f26886b5ba2f0ffd2d69586f5fa6482e7eab

Contents

achim	2
briggs_maccallum2003	2
caron2016	3
cormat	3
cormat.l	4
ex_2factors	4
ex_3factors_doub_unique	5
ex_4factors_corr	5
genr8	6
loadings	6
meek_bouchard	7
nest	8
pa	10

plot.nest	11
print.nest	11
shem	12
summary.nest	13

Index	14
--------------	-----------

achim	<i>A list of seven correlation matrices.</i>
-------	--

Description

A a list of seven correlation matrices. Given by Achim, A (personal communication).

Usage

achim

Format

A 12 by 12 correlation matrix

Source

<https://github.com/quantmeth>

briggs_maccallum2003	<i>A list of three correlation matrice.</i>
----------------------	---

Description

See Briggs, N. E., & MacCallum, R. C. (2003). Recovery of weak common factors by Maximum likelihood and ordinary least squares estimation. *Multivariate Behavioral Research*, 38(1), 25–56. [doi:10.1207/S15327906MBR3801_2](https://doi.org/10.1207/S15327906MBR3801_2)

Usage

briggs_maccallum2003

Format

A a list of three correlation matrices found in Briggs & MacCullum (2003).

Source

<https://github.com/quantmeth>

caron2016	<i>A list of six correlation matrices composed of nine variables with three factors.</i>
-----------	--

Description

See Caron, P.-O. (2016). A Monte Carlo examination of the broken-stick distribution to identify components to retain in principal component analysis. *Journal of Statistical Computation and Simulation*, 86(12), 2405-2410. doi:10.1080/00949655.2015.1112390

Usage

caron2016

Format

A list of six 9 x 9 correlation matrices found in Caron (2016).

Source

<https://github.com/quantmeth>

cormat	<i>A list containing 120 correlation matrices</i>
--------	---

Description

A list containing 120 24×24 correlation matrices (R) built to represent different factor structures. Details are found in the 'cormat.l' data.

Usage

cormat

Format

A a list of 120 correlation matrices

Source

<https://github.com/quantmeth>

cormat.l	<i>A list containing 120 lists of correlation matrices and their underlying characteristics</i>
----------	---

Description

A list containing 120 lists of 24×24 correlation matrices (R) built to represent different factor structures. Different levels of loadings (delta, .4, .5, .6, .7, .8), correlation between factors (corrfact, .0, .1, .2 .3), and number of factors (nfactors, 1:8) are used. The list contained matrice (R), and their underlying characteristics (delta, corrfact, and nfactors).

Usage

```
cormat.l
```

Format

A list containing 120 matrices

Source

<https://github.com/quantmeth>

ex_2factors	<i>A correlation matrix composed of 2 factors.</i>
-------------	--

Description

A correlation matrix composed of 10 items based on 2 factors with 5 variables each and loadings equals to .80.

Usage

```
ex_2factors
```

Format

A 10 by 10 correlation matrix

Source

<https://github.com/quantmeth>

ex_3factors_doub_unique

A correlation matrix composed of two factors, a double factor and a unique variable.

Description

A correlation matrix composed of 10 items based on two main factors among which there is two cross-loadings. There is also a double factors and an unique variable. Given by Achim, A. (personal communication).

Usage

ex_3factors_doub_unique

Format

A 10 by 10 correlation matrix

Source

<https://github.com/quantmeth>

ex_4factors_corr

A correlation matrix composed of 4 correlated factors.

Description

A correlation matrix composed of 12 items based on 4 factors with 3 variables each. Loadings equals to .9, .9, and .3. Factors 1 and 2, and factors 3 and 4 are correlated at .7. Given by Achim, A (personal communication).

Usage

ex_4factors_corr

Format

A 12 by 12 correlation matrix

Source

<https://github.com/quantmeth>

genr8 *Simplify the the generation from a Multivariate Normal Distributions*

Description

Speeds up the use of MASS::mvrnorm

Usage

```
genr8(n = 1, R = diag(10), mean = rep(0, ncol(R)), ...)
```

Arguments

n the number of samples required.

R a positive-definite symmetric matrix specifying the covariance matrix of the variables.

mean an optimal vector giving the means of the variables. Default is 0.

... Arguments for MASS::mvrnorm(), such as tol, empirical, and EISPACK.

Value

A data frame of size n by ncol(R).

Examples

```
set.seed(19)
R <- caron2016$mat1
mydata <- genr8(n = nrow(R)+1, R = R, empirical = TRUE)
round(mydata, 2)
round(cov(mydata), 2)
```

loadings *Print Loadings in NEST*

Description

Print Loadings in NEST

Usage

```
loadings(x, nfactors = x$nfactors, method = x$method, ...)
```

Arguments

<code>x</code>	An object of class "nest".
<code>nfactors</code>	The number of factors to retains.
<code>method</code>	A method used to compute loadings and uniquenesses.
<code>...</code>	Further arguments to methods in "nest" or the <code>stats::loadings</code> function.

Value

A $p \times k$ matrix containing loadings where p is the number of variables and k is the number of factors (`nfactors`).

Note

See `stats::loadings` for the original documentation.

Examples

```
results <- nest(ex_2factors, n = 100)
loadings(results)
```

meek_bouchard	<i>A correlation matrix given by Meek-Bouchard.</i>
---------------	---

Description

An empirical correlation matrix composed of 44 items given by Meek-Bouchard, C. (personal communication).

Usage

```
meek_bouchard
```

Format

A 44 by 44 correlation matrix

Source

<https://github.com/quantmeth>

nest *Nest Eigenvalue Sufficiency Test (NEST)*

Description

nest is used to identify the number of factors to retain in exploratory factor analysis.

Usage

```
nest(
  data,
  n = NULL,
  nrep = 1000,
  alpha = 0.05,
  max.fact = ncol(data),
  method = "ml",
  ...
)
```

Arguments

data	A data frame, a numeric matrix, covariance matrix or correlation matrix from which to determine the number of factors.
n	The number of cases (subjects, participants, or units) if a covariance matrix is supplied in data.
nrep	The number of replications to simulate. Default is 1000.
alpha	A vector of type I error rates or $(1-\alpha)*100\%$ confidence intervals. Default is .05.
max.fact	An optional maximum number of factor to extract. Default is <code>max.fact = ncol(data)</code> .
method	A method used to compute loadings and uniquenesses. Four methods are implemented in Rnest : maximum likelihood method = "ml" (default), regularized common factor analysis method = "rcfa", minimum rank factor analysis method = "mrfa", and principal axis factoring method = "paf". See details for custom methods.
...	Arguments for method that can be supplied. See details.

Details

The Next Eigenvalues Sufficiency Test (NEST) is an extension of parallel analysis by adding a sequential hypothesis testing procedure for every $k = 1, \dots, p$ factor until the hypothesis is not rejected.

At $k = 1$, NEST and parallel analysis are identical. Both use an Identity matrix as the correlation matrix. Once the first hypothesis is rejected, NEST uses a correlation matrix based on the loadings and uniquenesses of the k^{th} factorial structure. NEST then resamples the eigenvalues of this new correlation matrix. NEST stops when the $\$k_1^2\$$ eigenvalues is within the confidence interval.

There is two method already implemented in `nest` to extract loadings and uniquenesses: maximum likelihood ("ml"; default), principal axis factoring ("paf"), and minimum rank factor analysis ("mrfa"). The functions use as arguments: `covmat`, `n`, `factors`, and ... (supplementary arguments passed by `nest`). They return loadings and uniquenesses. Any other user-defined functions can be used as long as it is programmed likewise.

Value

`nest` returns an object of class `nest`. The functions `summary` and `plot` are used to obtain and show a summary of the results.

An object of class `nest` is a list containing the following components:

- `nfactors` - The number of factors to retains (one by `alpha`).
- `cor` - The supplied correlation matrix.
- `n` - The number of cases (subjects, participants, or units).
- `values` - The eigenvalues of the supplied correlation matrix.
- `alpha` - The type I error rate.
- `method` - The method used to compute loadings and uniquenesses.
- `nrep` - The number of replications used.
- `prob` - Probabilities of each factor.
- `Eig` - A list of simulated eigenvalues.

Generic function

`plot.nest` Scree plot of the eigenvalues and the simulated confidence intervals for `alpha`.

`loadings` Extract loadings. It does not overwrite `stat::loadings`.

Author(s)

P.-O. Caron

References

Achim, A. (2017). Testing the number of required dimensions in exploratory factor analysis. *The Quantitative Methods for Psychology*, 13(1), 64-74. [doi:10.20982/tqmp.13.1.p064](https://doi.org/10.20982/tqmp.13.1.p064)

Examples

```
nest(ex_2factors, n = 100)
nest(mtcars)
```

pa *Parallel analysis*

Description

Parallel analysis

Usage

```
pa(  
  data = NULL,  
  n = NULL,  
  p = NULL,  
  nrep = 1000,  
  alpha = 0.05,  
  crit = NULL,  
  ...  
)
```

Arguments

data	data.frame.
n	number of subjects.
p	number of variables.
nrep	number of replications.
alpha	type I error rate.
crit	Critical values to compare the eigenvalues.
...	Other arguments

Value

nfactors (if data is supplied) and sampled eigenvalues

Examples

```
pa(ex_2factors, n = 42)  
E <- pa(n = 10, p = 2, nrep = 5)
```

plot.nest	<i>Print results of NEST</i>
-----------	------------------------------

Description

Scree plot of the eigenvalues and the $(1-\alpha)*100\%$ confidence intervals derived from the re-sampled eigenvalues supplied to nest.

Usage

```
## S3 method for class 'nest'  
plot(x, pa = FALSE, y, ...)
```

Arguments

x	An object of class "nest".
pa	Show results of Parallel Analysis.
y	Further arguments for other methods, ignored for "nest".
...	Further arguments for other methods, ignored for "nest".

Value

A ggplot output.

Note

This function is more interesting with many alpha values.

Examples

```
results <- nest(ex_2factors, n = 100, alpha = c(.01, .05, .01))  
plot(results)  
# Return the data used to produce the plot  
df <- plot(results)$data
```

print.nest	<i>Print results of NEST</i>
------------	------------------------------

Description

Print the number of factors to retain according to confidence levels.

Usage

```
## S3 method for class 'nest'  
print(x, ...)
```

Arguments

`x` An object of class "nest".
`...` Further arguments for other methods, ignored for "nest".

Value

No return value, called for side effects.

Examples

```
results <- nest(ex_2factors, n = 100)
print(results)
```

shem

Split-Half Eigenvector Matching (SHEM)

Description

shem estimates the number of principal components via Split-Half Eigenvector Matching (SHEM).

Usage

```
shem(data, nIts = 30)
```

Arguments

`data` A data frame, a numeric matrix, covariance matrix or correlation matrix from which to determine the number of factors.
`nIts` Number of iterations.

Value

shem returns a list containing the number of components, `nfactors`, whether the additional step in case of zero true latent components was carried, `zeroComponents`, the eigenvalues and the eigenvectors of the solution.

References

Galdwin, T. E. (2023) Estimating the number of principal components via Split-Half Eigenvector Matching (SHEM). *MethodsX*, 11, 102286. doi:10.1016/j.mex.2023.102286

Examples

```
jd <- genr8(n = 404, R = ex_4factors_corr)
shem(jd)
```

summary.nest	<i>Summary results of NEST</i>
--------------	--------------------------------

Description

summary method for class "nest".

Usage

```
## S3 method for class 'nest'  
summary(object, ...)
```

Arguments

object	An object of class "nest".
...	Further arguments for other methods, ignored for "nest".

Value

No return value, called for side effects.

Examples

```
results <- nest(ex_2factors, n = 100)  
summary(results)
```

Index

* datasets

- achim, 2
- briggs_maccallum2003, 2
- caron2016, 3
- cormat, 3
- cormat.l, 4
- ex_2factors, 4
- ex_3factors_doub_unique, 5
- ex_4factors_corr, 5
- meek_bouchard, 7

achim, 2

briggs_maccallum2003, 2

caron2016, 3

cormat, 3

cormat.l, 4

ex_2factors, 4

ex_3factors_doub_unique, 5

ex_4factors_corr, 5

genr8, 6

loadings, 6

meek_bouchard, 7

nest, 8

pa, 10

plot.nest, 11

print.nest, 11

shem, 12

summary.nest, 13