Package ‘fspe’

March 28, 2023

Type Package

Title Estimating the Number of Factors in EFA with Out-of-Sample Prediction Errors

Version 0.1.2

BugReports https://github.com/jmbh/fspe/issues

Depends R (>= 3.5.0)


License GPL-2

Encoding UTF-8

LazyData true

Imports psych, corpcor

NeedsCompilation no

Author Jonas Haslbeck [aut, cre],
Riet van Bork [ctb]

Maintainer Jonas Haslbeck <jonashaslbeck@gmail.com>

Repository CRAN

Date/Publication 2023-03-28 09:40:02 UTC

R topics documented:

  fspe .......................................................... 2
  holzinger19 ................................................. 3

Index 4
Estimating number of factors with out-of-sample prediction error

Description

Estimates the number of factors in Exploratory Factor Analysis (EFA) by approximating the out-of-sample prediction errors using a cross-validation scheme.

Usage

fspe(data, maxK, nfold = 10, rep = 1, method = "PE", rotate = "oblimin", pbar = TRUE, ...)

Arguments

data: A n x p matrix or data.frame.
maxK: The largest number of factors considered. For example, if maxK = 8, the factors 1, 2, ..., 8 are considered.
nfold: The number of folds in the cross-validation scheme. Defaults to nfold = 10.
rep: The number of repetitions of the cross-validation scheme. The repetitions are aggregated by averaging the prediction errors for each number of factor and subsequently selecting the number of factors that minimizes this average. Defaults to rep = 1.
method: If method = "PE", the factor estimation is based on the prediction error on the variables; If method = "Cov" the error is computed on the covariance matrix. Defaults to method = "PE" since it has performed better in simulations (see Haslbeck & van Bork, 2021 below.)
rotate: The rotation being used in the underlying factor analysis models. The options are the same as in the fa() function in the psych package, which is called by fpse. Defaults to rotate = "oblimin".
pbar: If pbar = TRUE, a progress bar is shown.
...

Details

The function returns:

Value

nfactor: An integer indicating the estimated number of factors.
nfactor: A 4-dimensional (for the PE method: variables, candidate models, folds, repetitions) or a 3-dimensional (for the Cov method: candidate models, folds, repetitions) array containing all prediction errors.
holzinger19

Author(s)
Jonas Haslbeck <jonashaslbeck@gmail.com>

References

Examples

```r
data(holzinger19)
fspe_out <- fspe(holzinger19,
    maxK = 10,
    nfold = 10,
    method = "PE")

fspe_out$nfactor # estimated factors = 4

# Plot OoS PE path
plot(fspe_out$PEs, xlab="No. Factors", ylab="OoS PE")
```

holzinger19

**Psychometric example data set by Holzinger & Swineford (1936)**

Description
19 ability measurements from the classic psychometrics data set of Holzinger and Swineford (1939). The original data set contains 24 variables, however, we excluded the last four to obtain an example data set for four factors (see Harman, 1967). We took this dataset from the R-package psychTools (Revelle & Revelle, 2020).

Usage

data(holzinger19)

References
Index

fspe, 2
holzinger19, 3