

Package ‘fwdmsa’

February 19, 2015

Version 0.2

Date 2011-07-26

Title Forward search for Mokken scale analysis

Author Wobbe P. Zijlstra <w.p.zijlstra@uvt.nl>

Maintainer Wobbe P. Zijlstra <w.p.zijlstra@uvt.nl>

Depends R (>= 2.10.0)

Description fwdmsa performs the Forward Search for Mokken scale analysis. It detects outliers, it produces several types of diagnostic plots.

License GPL (>= 2)

URL <http://www.tilburguniversity.edu/webwijs/show/?uid=w.p.zijlstra>

Repository CRAN

Date/Publication 2011-07-27 09:05:24

NeedsCompilation no

R topics documented:

fwdmsa-package	2
acs	4
fs.MSA	6
fs.MSA.n1	8
plot.fs.class	10
plot.fs.n1.class	13
Index	15

fwdmsa-package	<i>Robust Mokken Scale Analysis by Means of the Forward Search Algorithm for Outlier Detection</i>
----------------	--

Description

The package conducts the Forward Search on test and questionnaire data, and shows forward plots for the detection of outliers.

Details

Package: fwdmsa
Type: Package
Version: 0.2
Date: 2011-07-26
License: GPL Version 2 or later

The package includes the functions

<code>fs.MSA</code>	Computes the necessary input for forward plots
<code>plot.fs.class</code>	S3 method for forward plots
<code>fs.MSA.n1</code>	Computes $n1$
<code>plot.fs.n1.class</code>	S3 method for a plot showing graphically $n1$

and data set

`acs` Autonomy-Connectedness Scale

Thanks are due to L. Andries van der Ark for contributing R code, and Marrie Bekker and Marcel van Assen for providing the data set.

Author(s)

Wobbe P. Zijlstra Maintainer: Wobbe P. Zijlstra <w.p.zijlstra@uvt.nl>.

References

Bekker M. H. J., and Van Assen, M. A. L. M. (2006). A short form of the autonomy scale: Properties of the autonomy-connectedness scale (ACS-30). *Journal of Personality Assessment*, 86, 51-60.

Van der Ark, L. A. (2007). Mokken scale analysis in **R**. *Journal of Statistical Software*. <http://www.jstatsoft.org>

Zijlstra, W. P., Van der Ark, L. A., and Sijtsma, K. (2011). Robust Mokken scale analysis by means of the forward search algorithm for outlier detection. *Multivariate Behavioral Research*, 46, 58-89.

Examples

```
## Not run:
## Analyses of Zijlstra et al. (2010).
## First Forward Search Analysis
  library(fwdmsa)
  data(acs.cont)

# Determining n1 = 292
# Takes approximately 40 minutes
fs1.1.n1 <- fs.MSA.n1(acs.cont, B=100)
n1 <- fs1.1.n1$n1

# Figure 2: Plot of number unique subsamples
plot(fs1.1.n1)

# Running the forward search
fs1.1 <- fs.MSA(acs.cont)

# Figure 3: Plot of objective function
plot(fs1.1, type="objective", observations=1:618, col="gray70", n0=TRUE, n1=fs.res.cont.n1$n1, xlim=c(0,650))
plot(fs1.1, type="objective", id.observation=619, col=1, lwd=2, lty=2, add=TRUE)
plot(fs1.1, type="objective", observations=589:618, lwd=2, add=TRUE)

# Figure 4: Gap plot
plot(fs1.1, type="gap", ylim=c(-10,12), n0=TRUE, n1=292)

# Figure 5: Follow-up plots
plot(fs1.1, type="followup", step=543:548, reference.step=543, n0=TRUE, n1=292)

## Remove influential observations from the data
acs.sus <- acs.cont[-(589:618),]

# Determining n1 = 296
fs1.2.n1 <- fs.MSA.n1(acs.sus, B=100)
n1 <- fs1.2.n1$n1

# Running the forward search
fs1.2 <- fs.MSA(acs.sus)

# Figure 6: Minexcl plot
plot(fs1.2, type="minexcl", n0=TRUE, n1=296, n2=TRUE)

# Figure 7: Plot of number of scales
plot(fs1.2, type="num.scale", n0=TRUE, n1=296, n2=TRUE)

# Figure 8: Item entry plot for the longest scale
plot(fs1.2, type="scale", id.scale=1, n0=TRUE, n1=296, n2=TRUE)
```

```

## Second Forward Search Analysis
# Remove bad items from the data
acs.min.core <- acs.cont[-(589:618),-c(3,7,8,11,13,16)]

# Determining n1 = 302
fs2.n1 <- fs.MSA.n1(acs.min.core, B=100)
n1 <- fs2.1.n1$n1

# Running the forward search
fs2 <- fs.MSA(acs.min.core)

# Figure 9: Plot of restsore regression of item 1 for steps 302 and 589
plot(fs2, type="restscore", step=302, items=1, lty=2, ylim=c(0,4), n0=TRUE, n1=302, n2=TRUE)
plot(fs2, type="restscore", step=589, items=1, lty=1, add=TRUE)

# Figure 10: Plot of estimated IRF of item 1
plot(fs2, type="IRF", items=1, n0=TRUE, n1=302, n2=TRUE)

# Figure 11: Plot of coefH
plot(fs2, type="coefH", n0=TRUE, n1=302, n2=TRUE, ylim=c(.1,.8))

## What if influential observations were not removed from the data
acs.cont.core <- acs.cont[-c(3,7,8,11,13,16)]
# Determining n1 = 347
fs3.n1 <- fs.MSA.n1(acs.cont.core, B=100)
n1 <- fs3.n1$n1

# Running the forward search
fs3 <- fs.MSA(acs.cont.core)

# Figure 12a: Plot of estimated IRF of item 1 with influential observations
plot(fs3, type="IRF", items=1, n0=TRUE, n1=347, n2=FALSE)

# Figure 12b: Plot of coefH with influential observations
plot(fs3, type="coefH", n0=TRUE, n1=347, n2=FALSE, ylim=c(.1,.8))

## End(Not run)

```

acs

Autonomy-Connectedness Scale

Description

acs are the scores of 588 students on 17 items from the Dutch version of the subscale “sensitivity to others” of the Autonomy-Connectedness Scale.

acs.cont are the acs data with additional 31 contaminants.

Usage

```

data(acs)
data(acs.cont)

```

Format

acs is a 588 by 17 matrix containing integers.

acs.cont is a 619 by 17 matrix containing integers.

Details

Each item has five ordered answer categories (0 = disagree, 1 = slightly disagree, 2 = agree nor disagree, 3 = slightly agree, 4 = agree). The items are shown in the table below. An “N” indicates a negatively worded item. The scores on the negatively worded items were reversed.

- | | | |
|-----|--|---|
| 1. | Usually I can dismiss another person’s misery from my mind | N |
| 2. | If I have things my own way against the will of others, I usually get very restless | |
| 3. | I hate detachment | |
| 4. | I am seldom occupied with the feelings and experiences of others | N |
| 5. | I easily put aside other people’s comments | N |
| 6. | I am rarely occupied with other people’s view of me | N |
| 7. | If I imagine myself having to say goodbye to a beloved person, I feel brokenhearted in advance | |
| 8. | I am seldom inclined to ask other people’s advice | N |
| 9. | I often go deeply into other people’s feelings | |
| 10. | I often wonder what other people think of me | |
| 11. | When I take important decisions about my life, I leave other people’s wishes and opinions out of consideration | N |
| 12. | I can hardly bear it when other people are angry with me | |
| 13. | Somebody else’s experiences leave a strong mark on my own moods | |
| 14. | I feel a strong need for other people’s advice and guidance | |
| 15. | If I do something that bothers other people, I can easily dismiss that from my mind | N |
| 16. | I often long for love and warmth | |
| 17. | I can easily back out of things that people who are important to me want me to do | N |

Source

Data were kindly made available by M. H. J. Bekker and M. A. L. M. van Assen from the Tilburg University.

References

Bekker M. H. J., and Van Assen, M. A. L. M. (2006). A short form of the autonomy scale: Properties of the autonomy-connectedness scale (ACS-30). *Journal of Personality Assessment*, 86, 51-60.

See Also

[fs.MSA](#), [plot.fs.class](#), [fs.MSA.n1](#), [plot.fs.n1.class](#)

Examples

```
data(acs)
data(acs.cont)
```

Description

Computes the necessary input for forward plots for Mokken scale analysis

Usage

```
fs.MSA(
  X,
  initial.subsample = "random",
  initial.subsample.size = default.initial.subsample.size,
  minsize = default.minsize,
  seed = default.seed,
  n.low = default.n.low,
  verbose = TRUE)
```

Arguments

<code>X</code>	Matrix or data frame of numeric data containing the responses of $nrow(X)$ respondents to $ncol(X)$ items. Each row is called an <i>observation</i> . Each item has $m + 1$ response options $0, \dots, m$. Other scores (e.g., $1, \dots, m + 1$), are converted to $0, \dots, m$. Missing values are not allowed.
<code>initial.subsample</code>	Default is "random". Otherwise, a vector containing the rows of <code>X</code> that make up the initial subsample.
<code>initial.subsample.size</code>	Integer giving the size of the initial subsample. Only effective if <code>initial.subsample="random"</code> . By default <code>initial.subsample.size</code> equals the minimum of the number of restscore groups over all items multiplied by the number of items. If <code>initial.subsample</code> is not "random", the size of the initial subsample equals the length of the vector provided at <code>initial.subsample</code> .
<code>minsize</code>	Integer giving the minimum size of a rest score group. By default <code>minsize = N/10</code> if $N \geq 500$; <code>minsize = N/5</code> if $250 \leq N < 500$; and <code>minsize = max(N/3, 50)</code> if $N < 250$
<code>seed</code>	Numeric; fixes the random number generation set .seed in order to control the initial subsample. Default is a randomly drawn value between 1 and 10000.
<code>n.low</code>	Numeric; $n2$ should be larger than <code>n.low</code> . Default <code>n.low = N/4</code> .
<code>verbose</code>	Logical, indicating whether the subsample size should be printed on the screen. If FALSE, no output is produced. The default is TRUE.

Details

Function `fs.MSA` computes the required input for forward plots (`plot.fs.class`). Therefore, its values should be assigned to an object. Function `fs.MSA` may take a long time for data if the number of items and/or observations is large. A large `initial.subsample.size` reduces the computation time but may affect the results.

Value

Object of class `fs.class` containing the required input for forward plots (`plot.fs.class`). Only few of the items are of direct interest:

`initial.subsample.size`. Shows n_0 .

`n2`. Shows n_2 .

`suspect`. Shows the suspect observations.

`order.objective.function`. Shows the observations in descending order of their objective function values, for each subsample. Hence, the element row 1 and column 189 is the most suspect observation given the subsample of size 189.

Author(s)

W. P. Zijlstra <w.p.zijlstra@uvt.nl>

References

Zijlstra, W. P., Van der Ark, L. A., and Sijtsma, K. (2011). Robust Mokken scale analysis by means of the forward search algorithm for outlier detection. *Multivariate Behavioral Research*, 46, 58-89.

Van der Ark, L. A. (2007). Mokken scale analysis in R. *Journal of Statistical Software*. <http://www.jstatsoft.org>

See Also

`fs.MSA.n1`, `plot.fs.class`, `plot.fs.n1.class`

Examples

```
# Retrieve data (588 observations)
data(acs)

# Run Forward Search for Mokken scale analysis starting with
# 550 observations in the initial subsample size to save time
fwdmsa.res <- fs.MSA(acs, initial.subsample.size=550)

# Plot the objective function
plot(fwdmsa.res, xlim = c(540,588))

# Plot the objective function for observations 1, 2, and 4
plot(fwdmsa.res, id.observation = c(1,2,4), add=TRUE, col=2, xlim = c(540,588))

# Gap plot for subsamples 570 through 588
plot(fwdmsa.res, type = "gap", ylim = c(0,4), xlim = c(570,588))
```

```

# Follow-up plots
plot(fwdmsa.res, type="followup", step=560:565, reference.step=560, xlim = c(540,588))

# Min-excl plot.
plot(fwdmsa.res, type = "minexcl", n2=TRUE, xlim=c(540,588))

# Plot of number of scales
plot(fwdmsa.res, type="num.scale", n2=TRUE, xlim=c(540,588))

# Item entry plot for the longest scale
plot(fwdmsa.res, type="scale", id.scale=1, n2=TRUE, xlim=c(540,588))

# Plot of estimated IRF of item 1
plot(fwdmsa.res, type="IRF", items=1, n2=TRUE, xlim=c(540,588))

# Plot of coefH
plot(fwdmsa.res, type="coefH", n2=TRUE, ylim=c(.1,.8), xlim=c(540,588))

```

fs.MSA.n1

*Determining n1 for the Forward Search for Mokken Scale Analysis***Description**

Computes $n1$ for the Forward Search for Mokken scale analysis

Usage

```

fs.MSA.n1(
  X,
  B,
  cutoff = default.cutoff,
  initial.subsample.size = default.initial.subsample.size,
  minsize = default.minsize,
  seed = default.seed,
  verbose = TRUE)

```

Arguments

<code>X</code>	Matrix or data frame of numeric data containing the responses of $nrow(X)$ respondents to $ncol(X)$ items. Each row is called an <i>observation</i> . Each item has $m + 1$ response options $0, \dots, m$. Other scores (e.g., $1, \dots, m + 1$), are converted to $0, \dots, m$. Missing values are not allowed.
<code>B</code>	Integer giving the number of Forward Searches with different initial subsamples.
<code>cutoff</code>	Integer, the first step of the Forward Search for which the number of unique subsamples is below the cutoff equals $n1$.
<code>initial.subsample.size</code>	Integer giving the size of the initial subsample. By default <code>initial.subsample.size</code> equals the minimum of the number of restscore groups over all items multiplied by the number of items.

minsize	Integer giving the minimum size of a rest score group. By default minsize = $N/10$ if $N \geq 500$; minsize = $N/5$ if $250 \leq N < 500$; and minsize = $\max(N/3, 50)$ if $N < 250$
seed	Numeric; fixes the random number generation set.seed in order to control the initial subsample. Default is a randomly drawn value between 1 and 10000.
verbose	Logical, indicating whether B should be printed on the screen. If FALSE, no output is produced. The default is TRUE.

Details

Function fs.MSA.n1 computes the required input for the forward plot ([plot.fs.n1.class](#)). Therefore, its values should be assigned to an object. B should at least be larger than cutoff, preferably $B \geq 100$. Large values of B may take much computation time.

Value

number.unique.subsamples. The number of unique subsamples at each step of the Forward Search.

n1. The first step for which the number of unique subsamples is below the cutoff.

Author(s)

W. P. Zijlstra <w.p.zijlstra@uvvt.nl>

References

Zijlstra, W. P., Van der Ark, L. A., and Sijtsma, K. (2011). Robust Mokken scale analysis by means of the forward search algorithm for outlier detection. *Multivariate Behavioral Research*, 46, 58-89.

See Also

[plot.fs.n1.class](#), [fs.MSA](#), [plot.fs.class](#)

Examples

```
## Not run:
## Retrieve data (588 observations)
data(acs)

# Determine n1 by running the Forward Search for Mokken scale analysis
# B=100 times
fwdmsa.res.n1 <- fs.MSA.n1(acs, B=100)

# Plot of number unique subsamples
plot(fwdmsa.res.n1)

## End(Not run)
```

plot.fs.class	<i>Plot fs.class objects</i>
---------------	------------------------------

Description

S3 Method to plot objects of class fs.class. Graphic display of the robust Mokken scale analysis by means of the Forward Search. Seven different plots can be obtained.

Usage

```
## S3 method for class 'fs.class'
plot(
  x,
  type = "objective",
  observations = all.observations,
  id.observation = FALSE,
  items = all.items,
  id.item = FALSE,
  step = default.step,
  reference.step = default.reference.step,
  id.scale = default.scale,
  tukey.fences = TRUE,
  add = FALSE,
  n0 = FALSE,
  n1 = FALSE,
  n2 = FALSE,
  lower.c = default.lower.c,
  col = default.col,
  lwd = default.lwd,
  lty = default.lty,
  ylim = default.ylim,
  xlim = default.xlim,
  ...)
```

Arguments

x	Object of class fs.class produced by fs.MSA.
type	Type of forward plot: "objective" (default): Forward plot showing the objective function for an observation over all subsample sizes. "minexcl": Forward plot of the minimum objective function values of the observations excluded from the sample. "maxincl": Forward plot of the maximum objective function values of the observations included in the sample. "gap": Forward plot of the minexcl-maxincl. "coefH": Forward plot of Loevinger's scalability coefficient for items (H_j) and the test (H).

	"restscore": One graph for each item plotting the estimated expected item response functions.
	"IRF": Forward plot of the estimated expected item response functions.
	"followup": Forward plot of the observations entering and leaving the subsample at a specified step. The identity of the observations entering or leaving the subsample is provided in the plot.
	"scale": Forward plot showing whether the selected items belong to a scale.
	"num.scale": Forward plot of the number of scales found by the AISP.
observations	Vector containing the observations to show. The default uses all observations. Relevant for type="objective".
id.observation	Vector containing the observation(s) for which the results are plotted in a different color and the identity of the observation(s) is added to the plot. Relevant for type="objective".
items	Vector containing the items for which the results are plotted. Default the results for all items are depicted. Relevant for type="coefH", type="restscore", type="IRF", and type="scale".
id.item	Logical, if TRUE the identity of the items are added to the plot. Relevant for type="coefH".
step	Single number or vector containing the subsample size. Default is sample size N . For type="restscore" step is a single number. For type="individual" step may be a vector. Relevant for type="restscore" and type="individual".
reference.step	Single number containing the subsample size. Default is step-1. Relevant for type="individual".
id.scale	Numeric indicating which scale to show; id.scale=0 indicates unscalable items, id.scale=1 indicates the longest scale, id.scale=2 indicates the next longest scale, etc. Default shows all scales. Relevant for type="scale".
tukey.fences	Logical, if TRUE Tukey's Fences ($Q3 + 1.5 * IQR$ and $Q3 + 3 * IQR$) are plotted. Relevant for type="minexcl".
add	Logical, if TRUE, the plot is added to the current plot. The default is FALSE. Relevant for type="objective", type="minexcl", type="maxincl", type="gap", type="coefH", and type="restscore".
n0	Logical, indicating whether $n0$ should be added to the plots. Default is FALSE
n1	Default is FALSE otherwise a single number should be given
n2	Logical, indicating whether $n2$ should be added to the plots. Default is FALSE
lower.c	Numeric scaling criterium; $0 \leq \text{lowerbound} < 1$. The default is 0.3.
col	Colour to be used for the lines.
lwd	Line width
lty	Line type
ylim, xlim	Range of x and y values with sensible defaults.
...	Other arguments are ignored.

Details

add=TRUE for type="restscore" can only be used for one item (e.g., items=1).

Value

Returns a graph.

Author(s)

W. P. Zijlstra <w.p.zijlstra@uvvt.nl>

References

Zijlstra, W. P., Van der Ark, L. A., and Sijtsma, K. (2011). Robust Mokken scale analysis by means of the forward search algorithm for outlier detection. *Multivariate Behavioral Research*, 46, 58-89.

See Also

[fs.MSA](#), [fs.MSA.n1](#), [plot.fs.n1.class](#)

Examples

```
# Retrieve data (588 observations)
data(acs)

# Run Forward Search for Mokken scale analysis starting with
# 550 observations in the initial subsample size to save time
fwdmsa.res <- fs.MSA(acs, initial.subsample.size=550)

# Plot the objective function
plot(fwdmsa.res, xlim = c(540,588))

# Plot the objective function for observations 1, 2, and 4
plot(fwdmsa.res, id.observation = c(1,2,4), add=TRUE, col=2, xlim = c(540,588))

# Gap plot for subsamples 570 through 588
plot(fwdmsa.res, type = "gap", ylim = c(0,4), xlim = c(570,588))

# Follow-up plots
plot(fwdmsa.res, type="followup", step=560:565, reference.step=560, xlim = c(540,588))

# Min-excl plot.
plot(fwdmsa.res, type = "minexcl", n2=TRUE, xlim=c(540,588))

# Plot of number of scales
plot(fwdmsa.res, type="num.scale", n2=TRUE, xlim=c(540,588))

# Item entry plot for the longest scale
plot(fwdmsa.res, type="scale", id.scale=1, n2=TRUE, xlim=c(540,588))

# Plot of estimated IRF of item 1
```

```

plot(fwdmsa.res, type="IRF", items=1, n2=TRUE, xlim=c(540,588))

# Plot of coefH
plot(fwdmsa.res, type="coefH", n2=TRUE, ylim=c(.1,.8), xlim=c(540,588))

```

plot.fs.n1.class *Plot fs.n1.class objects*

Description

S3 Method to plot objects of class fs.n1.class. Graphic display of the number of unique subsamples produced by multiple Forward Searches.

Usage

```

## S3 method for class 'fs.n1.class'
plot(
  x,
  cutoff=default.cutoff,
  n1.type=default.type,
  lwd=default.lwd,
  lty=default.lty,
  col=default.col,
  ylim = default.ylim,
  xlim = default.xlim,
  ...)

```

Arguments

x	Object of class fs.n1.class produced by fs.MSA.n1 .
cutoff	Default is the cutoff used for determining <i>n1</i> in fs.MSA.n1 .
n1.type	Default is "unique" which shows number of unique subsamples; "major" shows the number of subsamples <i>not</i> belonging to largest unique subsample; "both" plots both unique and major.
lwd	Line width
lty	Line type
col	Colour to be used for the lines.
ylim, xlim	Range of x and y values with sensible defaults.
...	Other arguments are ignored.

Value

Returns a graph.

Author(s)

W. P. Zijlstra <w.p.zijlstra@uvvt.nl>

References

Zijlstra, W. P., Van der Ark, L. A., and Sijtsma, K. (2011). Robust Mokken scale analysis by means of the forward search algorithm for outlier detection. *Multivariate Behavioral Research*, 46, 58-89.

See Also

[fs.MSA.n1](#), [fs.MSA](#), [plot.fs.class](#)

Examples

```
## Not run:
## Retrieve data (588 observations)
data(acs)

# Determine n1 by running the Forward Search for Mokken scale analysis
# B=100 times
fwdmsa.res.n1 <- fs.MSA.n1(acs, B=100)

# Plot of number unique subsamples
plot(fwdmsa.res.n1)

# Plot of number of subsamples not belonging to largest unique subsample (majority)
plot(fwdmsa.res.n1, n1.type="major")

## End(Not run)
```

Index

*Topic **aplot**

plot.fs.class, 10

plot.fs.n1.class, 13

*Topic **datasets**

acs, 4

*Topic **package**

fwdmsa-package, 2

*Topic **univar**

fs.MSA, 6

fs.MSA.n1, 8

acs, 2, 4

fs.MSA, 2, 5, 6, 9, 10, 12, 14

fs.MSA.n1, 2, 5, 7, 8, 12–14

fwdmsa (fwdmsa-package), 2

fwdmsa-package, 2

plot.fs.class, 2, 5, 7, 9, 10, 14

plot.fs.n1.class, 2, 5, 7, 9, 12, 13