Package 'CFM'

October 22, 2025

Type Package

Title Analyzing Censored Factor Models

Version 0.6.0

Description Provides generation and estimation of censored factor models for high-dimensional data with censored errors (normal, t, logistic). Includes Sparse Orthogonal Principal Components (SOPC), and evaluation metrics. Based on Guo G. (2023) <doi:10.1007/s00180-022-01270-z>.

License MIT + file LICENSE

Encoding UTF-8

Language en-US

Depends R (>= 3.5.0)

Imports stats, MASS, psych, matrixcalc, crch

Suggests testthat (>= 3.0.0), ggplot2

NeedsCompilation no

RoxygenNote 7.3.3

Author Guangbao Guo [aut, cre],
Tong Meng [aut]

Maintainer Guangbao Guo <ggb11111111@163.com>

Repository CRAN

Date/Publication 2025-10-22 18:50:10 UTC

Contents

`censored_factor_models` *Basic censored-factor data simulator*

Description

Generates multivariate data that follow a latent factor structure with censored errors (Normal, Student-t or Logistic).

Usage

```
censored_factor_models(
  n,
  p,
  m,
  distribution = c("normal", "t", "logistic"),
  df = NULL,
  seed = NULL
)
```

Arguments

`n`	Sample size (> 0).
`p`	Number of observed variables (> 0).
`m`	Number of latent factors (< p).
`distribution`	Error distribution: "normal" (default), "t", "logistic".
`df`	Degrees of freedom when distribution = "t".
`seed`	Optional random seed.

Value

A list with components:

`data`	numeric n × p matrix of observations
`loadings`	p × m factor loadings matrix
`uniqueness`	p × p diagonal uniqueness matrix
`KMO`	KMO measure of sampling adequacy
`Bartlett_p`	p-value of Bartlett's test
`distribution`	error distribution used
`seed`	random seed

Examples

```
set.seed(2025)
obj <- censored_factor_models(200, 6, 2)
psych::KMO(obj$data)
```

CFM *Censored Factor Models Data Generation*

Description

Generate multivariate data that follow a latent factor structure with censoring errors drawn from Normal, Student-t or Logistic distributions. Convenience wrapper around `rcnorm`, `rct`, and `rclogis`.

Usage

```
CFM(n, p, m, cens.dist = c("normal", "t", "logistic"), df = 5, seed = NULL)
```

Arguments

`n`	sample size ($n \times 1$ observations).
`p`	number of manifest variables.
`m`	number of latent factors.
`cens.dist`	censoring error distribution: `"normal"`, `"t"`, or `"logistic"`.
`df`	degrees of freedom when `cens.dist = "t"`.
`seed`	optional random seed for reproducibility.

Value

A named list with components:

`data`	numeric $n \times p$ matrix of observations.
`F`	factor scores matrix ($n \times m$).
`A`	factor loadings matrix ($p \times m$).
`D`	unique variances diagonal matrix ($p \times p$).

Examples

```
set.seed(2025)
# Normal censoring
obj <- CFM(n = 200, p = 10, m = 3, cens.dist = "normal")
head(obj$data)

# t-censoring with 6 d.f.
obj <- CFM(n = 300, p = 12, m = 4, cens.dist = "t", df = 6)
psych::KMO(obj$data)
```

FanPC.CFM *Censored Factor Analysis via Principal Component (FanPC, pure R)*

Description

Censored Factor Analysis via Principal Component (FanPC, pure R)

Usage

```
FanPC.CFM(
  data,
  m,
  A = NULL,
  D = NULL,
  p = NULL,
  cens.dist = c("normal", "t", "logistic"),
  df = NULL,
  cens.method = c("winsorise", "em"),
  cens_prop = 0.01,
  surv.obj = NULL,
  ctrl = NULL,
  verbose = NULL
)
```

Arguments

data	Numeric matrix or data frame of dimension $n \times p$.
m	Number of factors (< p).
A	Optional true loading matrix, used only for error calculation.
D	Optional true unique-variance diagonal matrix, used only for error calculation.
p	Number of variables (deprecated; detected automatically).
cens.dist	Error distribution, reserved for future use.
df	Degrees of freedom, reserved for future use.
cens.method	Censoring handling method; currently only "winsorise" is implemented. Defaults to "winsorise".
cens_prop	Winsorisation proportion, default 0.01.
surv.obj	Reserved for future use.
ctrl	Reserved for future use.
verbose	Reserved for future use.

Value

AF Estimated loading matrix, p × m.

DF Estimated unique-variance diagonal matrix, p × p.

MSESigmaA Mean squared error of loadings (if A is provided).

MSESigmaD Mean squared error of unique variances (if D is provided).

LSigmaA Relative error of loadings (if A is provided).

LSigmaD Relative error of unique variances (if D is provided).

Examples

```
library(CFM)
obj <- CFM(n = 500, p = 10, m = 2, cens.dist = "normal")
res <- FanPC.CFM(obj$data, m = 2, A = obj$A, D = obj$D, cens.method = "winsorise")
print(res$MSESigmaA)
```

PC2.CFM *PC2 for censored factor models (Top-2 principal components, pure R)*

Description

PC2 for censored factor models (Top-2 principal components, pure R)

Usage

```
PC2.CFM(
  data,
  m,
  A = NULL,
  D = NULL,
  p = NULL,
  cens.dist = c("normal", "t", "logistic"),
  df = NULL,
  cens.method = c("winsorise", "em"),
  cens_prop = 0.01,
  surv.obj = NULL,
  ctrl = NULL,
  verbose = NULL
)
```

Arguments

`data`	Numeric matrix or data frame of dimension $n \times p$.
`m`	Number of factors (< p).
`A`	Optional true loading matrix, used only for error calculation.
`D`	Optional true unique-variance diagonal matrix, used only for error calculation.
`p`	Number of variables (deprecated; detected automatically).
`cens.dist`	Error distribution, reserved for future use.
`df`	Degrees of freedom, reserved for future use.
`cens.method`	Censoring handling method; currently only `"winsorise"` is implemented. Defaults to `"winsorise"`.
`cens_prop`	Winsorisation proportion, default 0.01.
`surv.obj`	Reserved for future use.
`ctrl`	Reserved for future use.
`verbose`	Reserved for future use.

Value

AF Estimated loading matrix, p × 2.

DF Estimated unique-variance diagonal matrix, p × p.

MSESigmaA Mean squared error of loadings (if A is provided).

MSESigmaD Mean squared error of unique variances (if D is provided).

LSigmaA Relative error of loadings (if A is provided).

LSigmaD Relative error of unique variances (if D is provided).

Examples

```
library(CFM)
obj <- CFM(n = 500, p = 12, m = 2, cens.dist = "normal")
res <- PC2.CFM(obj$data, A = obj$A, D = obj$D)
print(res$MSESigmaA)
```

PPC2.CFM *PPC2 for censored factor models (Top-2 principal components, pure R)*

Description

PPC2 for censored factor models (Top-2 principal components, pure R)

Usage

```
PPC2.CFM(
  data,
  m,
  A = NULL,
  D = NULL,
  p = NULL,
  cens.dist = c("normal", "t", "logistic"),
  df = NULL,
  cens.method = c("winsorise", "em"),
  cens_prop = 0.01,
  surv.obj = NULL,
  ctrl = NULL,
  verbose = NULL
)
```

Arguments

`data`	Numeric matrix or data frame of dimension $n \times p$.
`m`	Number of factors (< p).
`A`	Optional true loading matrix, used only for error calculation.
`D`	Optional true unique-variance diagonal matrix, used only for error calculation.
`p`	Number of variables (deprecated; detected automatically).
`cens.dist`	Error distribution, reserved for future use.
`df`	Degrees of freedom, reserved for future use.
`cens.method`	Censoring handling method; currently only `"winsorise"` is implemented. Defaults to `"winsorise"`.
`cens_prop`	Winsorisation proportion, default 0.01.
`surv.obj`	Reserved for future use.
`ctrl`	Reserved for future use.
`verbose`	Reserved for future use.

Value

AF Estimated loading matrix, p × 2.

DF Estimated unique-variance diagonal matrix, p × p.

MSESigmaA Mean squared error of loadings (if A is provided).

MSESigmaD Mean squared error of unique variances (if D is provided).

LSigmaA Relative error of loadings (if A is provided).

LSigmaD Relative error of unique variances (if D is provided).

Examples

```
library(CFM)
obj <- CFM(n = 500, p = 12, m = 2, cens.dist = "normal")
res <- PPC2.CFM(obj$data, A = obj$A, D = obj$D, cens.method = "winsorise")
print(res$MSESigmaA)
```

Index