Package 'hIRT'

October 13, 2022

Type Package

Title Hierarchical Item Response Theory Models

Version 0.3.0

Description Implementation of a class of hierarchical item response theory (IRT) models where both the mean and the variance of latent preferences (ability parameters) may depend on observed covariates. The current implementation includes both the two-parameter latent trait model for binary data and the graded response model for ordinal data. Both are fitted via the Expectation-Maximization (EM) algorithm. Asymptotic standard errors are derived from the observed information matrix.

Depends R (>= 3.4.0), stats

Imports pryr (>= 0.1.2), rms (>= 5.1-1), ltm (>= 1.1-1), Matrix (>= 1.2-10)

Suggests ggplot2 (>= 2.2.1), knitr, rmarkdown

License GPL (>= 3)

Encoding UTF-8

LazyData true

RoxygenNote 7.0.2

URL http://github.com/xiangzhou09/hIRT

BugReports http://github.com/xiangzhou09/hIRT

NeedsCompilation no

Author Xiang Zhou [aut, cre]

Maintainer Xiang Zhou <xiang_zhou@fas.harvard.edu>

Repository CRAN

Date/Publication 2020-03-26 17:10:02 UTC

R topics documented:

`coef_item` *Parameter Estimates from Hierarchical IRT Models.*

Description

Parameter estimates from either `hltm` or `hgrm` models. `code_item` reports estimates of item parameters. `coef_mean` reports results for the mean equation. `coef_var` reports results for the variance equation.

Usage

```
coef_item(x, by_item = TRUE, digits = 3)

coef_mean(x, digits = 3)

coef_var(x, digits = 3)
```

Arguments

`x`	An object of class `hIRT`
`by_item`	Logical. Should item parameters be stored item by item (if `TRUE`) or put together in a data frame (if `FALSE`)?
`digits`	The number of significant digits to use when printing

Value

Parameter estimates, standard errors, z values, and p values organized as a data frame (if `by_item = TRUE`) or a list (if `by_item = FALSE`).

Examples

```
y <- nes_econ2008[, -(1:3)]
x <- model.matrix( ~ party * educ, nes_econ2008)
z <- model.matrix( ~ party, nes_econ2008)
nes_m1 <- hgrm(y, x, z)
coef_item(nes_m1)
coef_mean(nes_m1)
coef_var(nes_m1)
```

---

hgrm *Fitting Hierarchical Graded Response Models (for Ordinal Responses)*

Description

hgrm fits a hierarchical graded response model in which both the mean and the variance of the latent preference (ability parameter) may depend on person-specific covariates (x and z). Specifically, the mean is specified as a linear combination of x and the log of the variance is specified as a linear combination of z. Nonresponses are treated as missing at random.

Usage

```
hgrm(
  y,
  x = NULL,
  z = NULL,
  constr = c("latent_scale", "items"),
  beta_set = 1L,
  sign_set = TRUE,
  init = c("naive", "glm", "irt"),
  control = list()
)
```

Arguments

y	A data frame or matrix of item responses.
x	An optional model matrix, including the intercept term, that predicts the mean of the latent preference. If not supplied, only the intercept term is included.
z	An optional model matrix, including the intercept term, that predicts the variance of the latent preference. If not supplied, only the intercept term is included.
constr	The type of constraints used to identify the model: "latent_scale", or "items". The default, "latent_scale" constrains the mean of latent preferences to zero and the geometric mean of prior variance to one; "items" places constraints on item parameters instead and sets the mean of item difficulty parameters to zero and the geometric mean of the discrimination parameters to one.

`beta_set` The index of the item for which the discrimination parameter is restricted to be positive (or negative). It may take any integer value from 1 to `ncol(y)`.

`sign_set` Logical. Should the discrimination parameter of the corresponding item (indexed by `beta_set`) be positive (if `TRUE`) or negative (if `FALSE`)?

`init` A character string indicating how item parameters are initialized. It can be "naive", "glm", or "irt".

`control` A list of control values

max_iter The maximum number of iterations of the EM algorithm. The default is 150.

eps Tolerance parameter used to determine convergence of the EM algorithm. Specifically, iterations continue until the Euclidean distance between β_n and β_{n-1} falls under `eps`, where β is the vector of item discrimination parameters. `eps`=1e-4 by default.

max_iter2 The maximum number of iterations of the conditional maximization procedures for updating γ and λ. The default is 15.

eps2 Tolerance parameter used to determine convergence of the conditional maximization procedures for updating γ and λ. Specifically, iterations continue until the Euclidean distance between two consecutive log likelihoods falls under `eps2`. `eps2`=1e-3 by default.

K Number of Gauss-Legendre quadrature points for the E-step. The default is 21.

C [-C, C] sets the range of integral in the E-step. `C`=3 by default.

Value

An object of class `hgrm`.

`coefficients` A data frame of parameter estimates, standard errors, z values and p values.

`scores` A data frame of EAP estimates of latent preferences and their approximate standard errors.

`vcov` Variance-covariance matrix of parameter estimates.

`log_Lik` The log-likelihood value at convergence.

`N` Number of units.

`J` Number of items.

`H` A vector denoting the number of response categories for each item.

`ylevels` A list showing the levels of the factorized response categories.

`p` The number of predictors for the mean equation.

`q` The number of predictors for the variance equation.

`control` List of control values.

`call` The matched call.

References

Zhou, Xiang. 2019. "Hierarchical Item Response Models for Analyzing Public Opinion." Political Analysis.

Examples

```
y <- nes_econ2008[, -(1:3)]
x <- model.matrix( ~ party * educ, nes_econ2008)
z <- model.matrix( ~ party, nes_econ2008)
nes_m1 <- hgrm(y, x, z)
nes_m1
```

---

hgrm2 *Hierarchical Graded Response Models with Known Item Parameters*

Description

hgrm2 fits a hierarchical graded response model where the item parameters are known and supplied by the user.

Usage

```
hgrm2(y, x = NULL, z = NULL, item_coefs, control = list())
```

Arguments

y	A data frame or matrix of item responses.
x	An optional model matrix, including the intercept term, that predicts the mean of the latent preference. If not supplied, only the intercept term is included.
z	An optional model matrix, including the intercept term, that predicts the variance of the latent preference. If not supplied, only the intercept term is included.
item_coefs	A list of known item parameters. The parameters of item j are given by the jth element, which should be a vector of length H_j, containing $H_j - 1$ item difficulty parameters (in descending order) and one item discrimination parameter.
control	A list of control values

max_iter The maximum number of iterations of the EM algorithm. The default is 150.

eps Tolerance parameter used to determine convergence of the EM algorithm. Specifically, iterations continue until the Euclidean distance between β_n and β_{n-1} falls under eps, where β is the vector of item discrimination parameters. eps=1e-4 by default.

max_iter2 The maximum number of iterations of the conditional maximization procedures for updating γ and λ. The default is 15.

eps2 Tolerance parameter used to determine convergence of the conditional maximization procedures for updating γ and λ. Specifically, iterations continue until the Euclidean distance between two consecutive log likelihoods falls under eps2. eps2=1e-3 by default.

K Number of Gauss-Legendre quadrature points for the E-step. The default is 21.

C [-C, C] sets the range of integral in the E-step. C=3 by default.

Value

An object of class hgrm.

coefficients	A data frame of parameter estimates, standard errors, z values and p values.
scores	A data frame of EAP estimates of latent preferences and their approximate standard errors.
vcov	Variance-covariance matrix of parameter estimates.
log_Lik	The log-likelihood value at convergence.
N	Number of units.
J	Number of items.
H	A vector denoting the number of response categories for each item.
ylevels	A list showing the levels of the factorized response categories.
p	The number of predictors for the mean equation.
q	The number of predictors for the variance equation.
control	List of control values.
call	The matched call.

Examples

```
y <- nes_econ2008[, -(1:3)]
x <- model.matrix( ~ party * educ, nes_econ2008)
z <- model.matrix( ~ party, nes_econ2008)

n <- nrow(nes_econ2008)
id_train <- sample.int(n, n/4)
id_test <- setdiff(1:n, id_train)

y_train <- y[id_train, ]
x_train <- x[id_train, ]
z_train <- z[id_train, ]

mod_train <- hgrm(y_train, x_train, z_train)

y_test <- y[id_test, ]
x_test <- x[id_test, ]
z_test <- z[id_test, ]

item_coefs <- lapply(coef_item(mod_train), `[[`, "Estimate")

model_test <- hgrm2(y_test, x_test, z_test, item_coefs = item_coefs)
```

`hltm` *Fitting Hierarchical Latent Trait Models (for Binary Responses).*

Description

`hltm` fits a hierarchical latent trait model in which both the mean and the variance of the latent preference (ability parameter) may depend on person-specific covariates (`x` and `z`). Specifically, the mean is specified as a linear combination of `x` and the log of the variance is specified as a linear combination of `z`.

Usage

```
hltm(
  y,
  x = NULL,
  z = NULL,
  constr = c("latent_scale", "items"),
  beta_set = 1L,
  sign_set = TRUE,
  init = c("naive", "glm", "irt"),
  control = list()
)
```

Arguments

`y`	A data frame or matrix of item responses.
`x`	An optional model matrix, including the intercept term, that predicts the mean of the latent preference. If not supplied, only the intercept term is included.
`z`	An optional model matrix, including the intercept term, that predicts the variance of the latent preference. If not supplied, only the intercept term is included.
`constr`	The type of constraints used to identify the model: "latent_scale", or "items". The default, "latent_scale" constrains the mean of latent preferences to zero and the geometric mean of prior variance to one; "items" places constraints on item parameters instead and sets the mean of item difficulty parameters to zero and the geometric mean of the discrimination parameters to one.
`beta_set`	The index of the item for which the discrimination parameter is restricted to be positive (or negative). It may take any integer value from 1 to `ncol(y)`.
`sign_set`	Logical. Should the discrimination parameter of the corresponding item (indexed by `beta_set`) be positive (if `TRUE`) or negative (if `FALSE`)?
`init`	A character string indicating how item parameters are initialized. It can be "naive", "glm", or "irt".
`control`	A list of control values

max_iter The maximum number of iterations of the EM algorithm. The default is 150.

eps Tolerance parameter used to determine convergence of the EM algorithm. Specifically, iterations continue until the Euclidean distance between β_n and β_{n-1} falls under `eps`, where β is the vector of item discrimination parameters. `eps`=1e-4 by default.

max_iter2 The maximum number of iterations of the conditional maximization procedures for updating γ and λ. The default is 15.

eps2 Tolerance parameter used to determine convergence of the conditional maximization procedures for updating γ and λ. Specifically, iterations continue until the Euclidean distance between two consecutive log likelihoods falls under `eps2`. `eps2`=1e-3 by default.

K Number of Gauss-Legendre quadrature points for the E-step. The default is 21.

C [-C, C] sets the range of integral in the E-step. `C`=3 by default.

Value

An object of class `hltm`.

`coefficients`	A data frame of parameter estimates, standard errors, z values and p values.
`scores`	A data frame of EAP estimates of latent preferences and their approximate standard errors.
`vcov`	Variance-covariance matrix of parameter estimates.
`log_Lik`	The log-likelihood value at convergence.
`N`	Number of units.
`J`	Number of items.
`H`	A vector denoting the number of response categories for each item.
`ylevels`	A list showing the levels of the factorized response categories.
`p`	The number of predictors for the mean equation.
`q`	The number of predictors for the variance equation.
`control`	List of control values.
`call`	The matched call.

References

Zhou, Xiang. 2019. "Hierarchical Item Response Models for Analyzing Public Opinion." Political Analysis.

Examples

```
y <- nes_econ2008[, -(1:3)]
x <- model.matrix( ~ party * educ, nes_econ2008)
z <- model.matrix( ~ party, nes_econ2008)

dichotomize <- function(x) findInterval(x, c(mean(x, na.rm = TRUE)))
y[] <- lapply(y, dichotomize)
nes_m1 <- hltm(y, x, z)
nes_m1
```

hltm2 *Hierarchical Latent Trait Models with Known Item Parameters.*

Description

hltm2 fits a hierarchical latent trait model where the item parameters are known and supplied by the user.

Usage

```
hltm2(y, x = NULL, z = NULL, item_coefs, control = list())
```

Arguments

y	A data frame or matrix of item responses.
x	An optional model matrix, including the intercept term, that predicts the mean of the latent preference. If not supplied, only the intercept term is included.
z	An optional model matrix, including the intercept term, that predicts the variance of the latent preference. If not supplied, only the intercept term is included.
item_coefs	A list of known item parameters. The parameters of item j are given by the jth element, which should be a vector of length 2, containing the item difficulty parameter and item discrimination parameter.
control	A list of control values

max_iter The maximum number of iterations of the EM algorithm. The default is 150.

eps Tolerance parameter used to determine convergence of the EM algorithm. Specifically, iterations continue until the Euclidean distance between β_n and β_{n-1} falls under eps, where β is the vector of item discrimination parameters. eps=1e-4 by default.

max_iter2 The maximum number of iterations of the conditional maximization procedures for updating γ and λ. The default is 15.

eps2 Tolerance parameter used to determine convergence of the conditional maximization procedures for updating γ and λ. Specifically, iterations continue until the Euclidean distance between two consecutive log likelihoods falls under eps2. eps2=1e-3 by default.

K Number of Gauss-Legendre quadrature points for the E-step. The default is 21.

C [-C, C] sets the range of integral in the E-step. C=3 by default.

Value

An object of class hltm.

coefficients	A data frame of parameter estimates, standard errors, z values and p values.

scores	A data frame of EAP estimates of latent preferences and their approximate standard errors.
vcov	Variance-covariance matrix of parameter estimates.
log_Lik	The log-likelihood value at convergence.
N	Number of units.
J	Number of items.
H	A vector denoting the number of response categories for each item.
ylevels	A list showing the levels of the factorized response categories.
p	The number of predictors for the mean equation.
q	The number of predictors for the variance equation.
control	List of control values.
call	The matched call.

Examples

```
y <- nes_econ2008[, -(1:3)]
x <- model.matrix( ~ party * educ, nes_econ2008)
z <- model.matrix( ~ party, nes_econ2008)
dichotomize <- function(x) findInterval(x, c(mean(x, na.rm = TRUE)))
y_bin <- y
y_bin[] <- lapply(y, dichotomize)

n <- nrow(nes_econ2008)
id_train <- sample.int(n, n/4)
id_test <- setdiff(1:n, id_train)

y_bin_train <- y_bin[id_train, ]
x_train <- x[id_train, ]
z_train <- z[id_train, ]

mod_train <- hltm(y_bin_train, x_train, z_train)

y_bin_test <- y_bin[id_test, ]
x_test <- x[id_test, ]
z_test <- z[id_test, ]

item_coefs <- lapply(coef_item(mod_train), `[[`, "Estimate")

model_test <- hltm2(y_bin_test, x_test, z_test, item_coefs = item_coefs)
```

---

latent_scores *Estimates of Latent Preferences/Abilities*

Description

EAP estimates of latent preferences for either hltm or hgrm models.

Usage

```
latent_scores(x, digits = 3)
```

Arguments

x	An object of class hIRT
digits	The number of significant digits to use when printing

Value

A data frame of EAP estimates of latent preferences and their approximate standard errors.

Examples

```
y <- nes_econ2008[, -(1:3)]
x <- model.matrix( ~ party * educ, nes_econ2008)
z <- model.matrix( ~ party, nes_econ2008)
nes_m1 <- hgrm(y, x, z)
pref <- latent_scores(nes_m1)
require(ggplot2)
ggplot(data = nes_econ2008) +
geom_density(aes(x = pref$post_mean, col = party))
```

---

nes_econ2008 *Public Attitudes on Economic Issues in ANES 2008*

Description

A dataset containing gender, party ID, education, and responses to 10 survey items on economic issues from the American National Election Studies, 2008.

Usage

```
nes_econ2008
```

Format

A data frame with 2268 rows and 13 variables:

gender gender. 1: male; 2: female

party party identification: Democrat, independent, or Republican

educ education. 1: high school or less; 2: some college or above

health_ins7 Support for government or private health insurance, 7 categories

jobs_guar7 Support for government guarantee jobs and income, 7 categories

gov_services7 Should government reduce or increase spending on services?, 7 categories

FS_poor3 Federal spending on the poor, 3 categories

FS_childcare3 Federal spending on child care, 3 categories

FS_crime3 Federal spending on crime, 3 categories

FS_publicschools3 Federal spending on public schools, 3 categories

FS_welfare3 Federal spending on welfare, 3 categories

FS_envir3 Federal spending on environment, 3 categories

FS_socsec3 Federal spending on Social Security, 3 categories

`print.hIRT` *Printing an object of class* `hIRT`

Description

Printing an object of class `hIRT`

Usage

```
## S3 method for class 'hIRT'
print(x, digits = 3, ...)
```

Arguments

`x`	An object of class `hIRT`
`digits`	The number of significant digits to use when printing
`...`	further arguments passed to `print`.

`summary.hIRT` *Summarizing Hierarchical Item Response Theory Models*

Description

Summarizing the fit of either `hltm` or `hgrm`.

Usage

```
## S3 method for class 'hIRT'
summary(object, by_item = FALSE, digits = 3, ...)

## S3 method for class 'summary_hIRT'
print(x, digits = 3, ...)
```

Arguments

object	An object of class hIRT.
by_item	Logical. Should item parameters be stored item by item (if TRUE) or put together in a data frame (if FALSE)?
digits	the number of significant digits to use when printing.
...	further arguments passed to print.
x	An object of class hIRT

Value

An object of class summary_hIRT.

call	The matched call.
model	Model fit statistics: Log likelihood, AIC, and BIC.
item_coefs	Item parameter estimates, standard errors, z values, and p values.
mean_coefs	Parameter estimates for the mean equation.
var_coefs	Parameter estimates for the variance equation.

Examples

```
y <- nes_econ2008[, -(1:3)]
x <- model.matrix( ~ party * educ, nes_econ2008)
z <- model.matrix( ~ party, nes_econ2008)
nes_m1 <- hgrm(y, x, z)
summary(nes_m1, by_item = TRUE)
```

Index