

# Package ‘hIRT’

June 13, 2019

**Type** Package

**Title** Hierarchical Item Response Theory Models

**Version** 0.2.0

**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)

**Suggests** ggplot2 (>= 2.2.1)

**License** GPL (>= 3)

**Encoding** UTF-8

**LazyData** true

**RoxygenNote** 6.1.0

**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** 2019-06-13 17:50:03 UTC

## R topics documented:

coef_item . . . . .	2
hgrm . . . . .	3

hltm . . . . .	4
latent_scores . . . . .	6
nes_econ2008 . . . . .	7
predict_hIRT . . . . .	7
print.hIRT . . . . .	8
summary.hIRT . . . . .	9

<b>Index</b>	<b>11</b>
--------------	-----------

---

coef_item	<i>Parameter Estimates from Hierarchical IRT Models.</i>
-----------	--

---

### Description

Parameter estimates from either `hltm` or `hgrm` models. `coef_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

<code>x</code>	An object of class <code>hIRT</code>
<code>by_item</code>	Logical. Should item parameters be stored item by item (if <code>TRUE</code> ) or put together in a data frame (if <code>FALSE</code> )?
<code>digits</code>	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)
```

## 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, z, beta_set = 1, sign_set = TRUE, control = list())
```

## Arguments

<code>y</code>	A data frame or matrix of item responses.
<code>x</code>	An optional model matrix, including the intercept term, that predicts the mean of the latent preference. If missing, only the intercept term is included.
<code>z</code>	An optional model matrix, including the intercept term, that predicts the variance of the latent preference. If missing, only the intercept term is included.
<code>beta_set</code>	The index of the item for which the discrimination parameter is restricted to be positive (or negative). It may take an integer value from 1 to <code>ncol(y)</code> .
<code>sign_set</code>	Logical. Should the discrimination parameter of the corresponding item (indexed by <code>beta_set</code> ) be positive (if TRUE) or negative (if FALSE)?
<code>control</code>	A list of control values

**max\_iter** The maximum number of iterations of the EM algorithm. Default 150.

**eps** Tolerance parameter used to determine convergence of the EM algorithm. Specifically, iterations continue until the Euclidean distance between  $\beta_n$  and  $\beta_{n-1}$  falls under `eps`, where  $\beta$  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  $\gamma$  and  $\lambda$ . Default 15.

**eps2** Tolerance parameter used to determine convergence of the conditional maximization procedures for updating  $\gamma$  and  $\lambda$ . 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. Default 21.

**C** [-C, C] sets the range of integral in the E-step. `C=5` 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)
print(nes_m1)
```

---

hltm

*Fitting Hierarchical Latent Trait Models (for Binary Data).*


---

**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, z, beta_set = 1, sign_set = TRUE, 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 missing, only the intercept term is included.
z	An optional model matrix, including the intercept term, that predicts the variance of the latent preference. If missing, only the intercept term is included.
beta_set	The index of the item for which the discrimination parameter is restricted to be positive (or negative). It may take an 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)?
control	A list of control values
	<b>max_iter</b> The maximum number of iterations of the EM algorithm. Default 150.
	<b>eps</b> Tolerance parameter used to determine convergence of the EM algorithm. Specifically, iterations continue until the Euclidean distance between $\beta_n$ and $\beta_{n-1}$ falls under eps, where $\beta$ is the vector of item discrimination parameters. eps=1e-4 by default.
	<b>max_iter2</b> The maximum number of iterations of the conditional maximization procedures for updating $\gamma$ and $\lambda$ . Default 15.
	<b>eps2</b> Tolerance parameter used to determine convergence of the conditional maximization procedures for updating $\gamma$ and $\lambda$ . Specifically, iterations continue until the Euclidean distance between two consecutive log likelihoods falls under eps2. eps2=1e-3 by default.
	<b>K</b> Number of Gauss-Legendre quadrature points for the E-step. Default 21.
	<b>C</b> [-C, C] sets the range of integral in the E-step. C=5 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)

# don't run
# nes_m1 <- hgrm(y, x, z)

dichotomize <- function(x) findInterval(x, c(mean(x, na.rm = TRUE)))
y_bin <- as.data.frame(lapply(y, dichotomize))
nes_m1 <- hltm(y_bin, x, z)
print(nes_m1)
```

---

latent_scores	<i>Estimates of Latent Preferences/Abilities</i>
---------------	--

---

## Description

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

## Usage

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

## Arguments

<code>x</code>	An object of class hIRT
<code>digits</code>	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$est, 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

predict\_hIRT

*Predicting from Hierarchical IRT Models***Description**

predict\_hIRT predicts latent preferences for new data  $(y, x, z)$  using a fitted hIRT object. Both posterior (given  $y, x, z$ ) and prior (given  $x, z$ ) estimates are returned.

**Usage**

```
predict_hIRT(object, y, x, z)
```

**Arguments**

object	An object of class hIRT.
y	A data frame or matrix of item responses.
x	Model matrix, including the intercept term, representing the covariates used to predict the means.
z	Model matrix, including the intercept term, representing the covariates used to predict the variances.

**Value**

A data frame.

post_mean	Posterior mean estimates of latent preferences, i.e., $E[\theta y, x, z]$
post_sd	Posterior standard deviation estimates of latent preferences, i.e., $\sqrt{V[\theta y, x, z]}$
prior_mean	Prior mean estimates of latent preferences, i.e., $E[\theta x]$
prior_sd	Prior standard deviation estimates of latent preferences, i.e., $\sqrt{V[\theta z]}$

**Examples**

```
n <- nrow(nes_econ2008)
n_train <- sample.int(n, 3*n/4)
n_test <- setdiff(1:n, n_train)

nes_econ2008_train <- nes_econ2008[n_train, , drop = FALSE]
nes_econ2008_test <- nes_econ2008[n_test, , drop = FALSE]

y <- nes_econ2008_train[, -(1:3)]
x <- model.matrix( ~ party * educ, nes_econ2008_train)
z <- model.matrix( ~ party, nes_econ2008_train)
mod_train <- hgrm(y, x, z)

y2 <- nes_econ2008_test[, -(1:3)]
x2 <- model.matrix( ~ party * educ, nes_econ2008_test)
z2 <- model.matrix( ~ party, nes_econ2008_test)
preds <- predict_hIRT(mod_train, y2, x2, z2)
summary(preds)
```

---

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 or from other methods

summary.hIRT

*Summarizing Hierarchical Item Response Theory Models***Description**

Summarizing the fit of either h1tm 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 or from other methods
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

## \*Topic **datasets**

nes\_econ2008, 7

coef\_item, 2

coef\_mean (coef\_item), 2

coef\_var (coef\_item), 2

hgrm, 3

hltm, 4

latent\_scores, 6

nes\_econ2008, 7

predict\_hIRT, 7

print.hIRT, 8

print.summary\_hIRT (summary.hIRT), 9

summary.hIRT, 9