Package 'lamle'

August 25, 2023

Encoding UTF-8

Type Package

Title Maximum Likelihood Estimation of Latent Variable Models

Version 0.3.1

Date 2023-08-24

Description Approximate marginal maximum likelihood estimation of multidimensional
latent variable models via adaptive quadrature or Laplace approximations to the
integrals in the likelihood function, as presented for confirmatory factor
analysis models in Jin, S., Noh, M., and Lee, Y. (2018)
<doi:10.1080/10705511.2017.1403287>, for item response theory models in
Andersson, B., and Xin, T. (2021) <doi:10.3102/1076998620945199>, and for
generalized linear latent variable models in Andersson, B., Jin, S., and
Zhang, M. (2023) <doi:10.1016/j.csda.2023.107710>. Models implemented include
the generalized partial credit model, the graded response model, and generalized
linear latent variable models for Poisson, negative-binomial and normal
distributions. Supports a combination of binary, ordinal, count and continuous
observed variables and multiple group models.

License GPL (>= 2)

Imports Rcpp (>= 1.0.1), mvtnorm, numDeriv, stats, fastGHQuad, methods

LinkingTo Rcpp, RcppArmadillo

NeedsCompilation yes

Author Björn Andersson [aut, cre] (<https://orcid.org/0000-0002-9007-2440>),
Shaobo Jin [aut] (<https://orcid.org/0000-0001-6538-3477>),
Maoxin Zhang [ctb]

Maintainer Björn Andersson <bjoern.h.andersson@gmail.com>

Repository CRAN

Date/Publication 2023-08-25 07:50:05 UTC

R topics documented:

---

lamle-package   *Maximum Likelihood Estimation of Latent Variable Models*

Description

Approximate marginal maximum likelihood estimation of multidimensional latent variable models via adaptive quadrature or Laplace approximations to the integrals in the likelihood function, as presented for confirmatory factor analysis models in Jin, S., Noh, M., and Lee, Y. (2018) <doi:10.1080/10705511.2017.1403287>, for item response theory models in Andersson, B., and Xin, T. (2021) <doi:10.3102/1076998620945199>, and for generalized linear latent variable models in Andersson, B., Jin, S., and Zhang, M. (2023) <doi:10.1016/j.csda.2023.107710>. Models implemented include the generalized partial credit model, the graded response model, and generalized linear latent variable models for Poisson, negative-binomial and normal distributions. Supports a combination of binary, ordinal, count and continuous observed variables and multiple group models.

Details

The DESCRIPTION file:

Package:	lamle
Encoding:	UTF-8
Type:	Package
Title:	Maximum Likelihood Estimation of Latent Variable Models
Version:	0.3.1
Date:	2023-08-24
Authors@R:	c(person(given = "Björn", family = "Andersson", role = c("aut", "cre"), email = "bjoern.h.andersson@gmail.co
Description:	Approximate marginal maximum likelihood estimation of multidimensional latent variable models via adaptiv
License:	GPL (>= 2)
Imports:	Rcpp (>= 1.0.1), mvtnorm, numDeriv, stats, fastGHQuad, methods
LinkingTo:	Rcpp, RcppArmadillo
Author:	Björn Andersson [aut, cre] (<https://orcid.org/0000-0002-9007-2440>), Shaobo Jin [aut] (<https://orcid.org/00
Maintainer:	Björn Andersson <bjoern.h.andersson@gmail.com>
Archs:	x64

Index of help topics:

DGP	Generation of Observed Data From a Generalized Linear Latent Variable Model
lamle	Estimation of Latent Variable Models with the Laplace Approximation or Adaptive Gauss-Hermite Quadrature
lamle-package	Maximum Likelihood Estimation of Latent Variable Models
lamle.compute	Compute Output from an Estimated Latent Variable Model
lamle.fit	Model Fit Statistics for an Estimated Latent Variable Model
lamle.plot	Plot Output from an Estimated Latent Variable Model
lamle.predict	Compute Latent Variable Estimates from an Estimated Latent Variable Model
lamle.sim	Generate Simulated Data from an Estimated Latent Variable Model
lamleout-class	Class "lamleout"

Author(s)

Björn Andersson, Shaobo Jin and Maoxin Zhang.

Maintainer: Björn Andersson <bjoern.h.andersson@gmail.com>

References

Andersson, B., Jin, S., and Zhang, M. (2023). Fast estimation of multiple group generalized linear latent variable models for categorical observed variables. *Computational Statistics and Data Analysis*, **182**, 1-12. <doi:10.1016/j.csda.2023.107710>

Andersson, B., and Xin, T. (2021). Estimation of latent regression item response theory models using a second-order Laplace approximation. *Journal of Educational and Behavioral Statistics*, **46(2)**, 244-265. <doi:10.3102/1076998620945199>

Huber, P., Ronchetti, E., and Victoria-Feser, M.P. (2004). Estimation of generalized linear latent variable models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, **66(4)**, 893-908. <doi:10.1111/j.1467-9868.2004.05627.x>

Jin, S., Noh, M., and Lee, Y. (2018). H-Likelihood Approach to Factor Analysis for Ordinal Data. *Structural Equation Modeling: A Multidisciplinary Journal*, **25(4)**, 530-540. <doi:10.1080/10705511.2017.1403287>

Shun, Z., and McCullagh, P. (1995). Laplace approximation of high dimensional integrals. *Journal of the Royal Statistical Society: Series B (Methodological)*, **57(4)**, 749-760. <doi:10.1111/j.2517-6161.1995.tb02060.x>

DGP	*Generation of Observed Data From a Generalized Linear Latent Variable Model*

Description

A function to generate simulated observed data based on a generalized linear latent variable model, supporting ordinal, count, and continuous observed variables.

Usage

```
DGP(a, b, modeltype, z)
```

Arguments

a	A matrix with number of rows equal to the number of observed variables and number of columns equal to the number of latent variables, containing the discrimination parameters for each item and dimension.
b	A list with length equal to the number of observed variables, containing intercept parameters and, if applicable, the scale parameter for each observed variable.
modeltype	A character vector with length equal to the number of observed variables to simulate, indicating which model should be used for each variable. Options are "GPCM", "GRM", "negbin", "normal", and "poisson".
z	The matrix of latent variable values, with rows indicating the individual and the columns indicating the dimension.

Value

A data matrix with the observed item responses.

Author(s)

Björn Andersson <bjoern.h.andersson@gmail.com>

References

Muraki, E. (1992). A Generalized Partial Credit Model: Application of an EM Algorithm. *Applied Psychological Measurement*, **16**, 159-176.

Examples

```
##### Load required package.
library(mvtnorm)

##### Item parameters
set.seed(123)
GPCMa <- runif(60, 0.8, 2)
GPCMb <- vector("list", 60)
for(i in 1:60) GPCMb[[i]] <- -c(runif(1, -3, -2),
                                runif(1, -1.5, -0.5),
                                runif(1, 0, 1),
                                runif(1, 1.5, 2.5))

##### Generate data from a two-dimensional independent-clusters model
```

```
mydata <- DGP(matrix(c(GPCMa[1:30], rep(0, 60), GPCMa[31:60]),
              ncol = 2), GPCMb, model = rep("GPCM", 60),
              rmvnorm(1000, c(0, 0), matrix(c(1, 0.6, 0.6, 1),
              ncol = 2)))
```

---

lamle   *Estimation of Latent Variable Models with the Laplace Approximation*
                         or Adaptive Gauss-Hermite Quadrature

---

Description

A function to estimate the parameters of generalized linear latent variables models using adaptive quadrature or Laplace approximations of the likelihood, with support for multiple groups and latent regression.

Usage

```
lamle(y, model, modeltype = rep("GRM", ncol(y)),
      link = rep("logit", ncol(y)), mi = NULL, X = NULL,
      group = NULL, parequal = NULL, parfix = NULL,
      groupequal = NULL, resdim = NULL, sw = NULL,
      method = "lap", accuracy = 2, tol = 1e-4, maxit = 500,
      optimizer = "BFGS", obsinfo = FALSE, adapt = TRUE,
      fullexp = TRUE, z.tol = 1e-8, maxdiff = 0.25,
      step.length = c(0.5, 1), first.step = 25,
      inithess = "NCW", startval = NULL, startmode = NULL,
      starteval = FALSE, checkgrad = FALSE, modeupdate = TRUE,
      verbose = FALSE, filtering = "advanced")
```

Arguments

y	The observed data matrix, with missing values coded as NA. Categorical data must be coded as successive integers starting from 1.
model	A matrix of NAs and 1s, with number of rows equal to the number of observed variables and number of columns equal to the number of latent variables, indicating which observed variables are related to which latent variable.
modeltype	A vector of strings with length equal to the number of observed variables, denoting which model to use for each observed variable. For ordinal data "GPCM" - generalized partial credit model, "GRM" - graded response model (default), for count data "negbin" - negative binomial, "poisson" - poisson, and for continuous data "normal" - normal.
link	A vector of strings with length equal to the number of observed variables, denoting which link function to use for each observed variable. Currently, this function is not active. The default link functions are logit (for "GPCM" and "GRM"), log (for "negbin" and "poisson"), and identity (for "normal").

mi	A vector with the number of categories of each categorical observed variable, where the default is to compute the number of categories automatically from the observed categorical variables.
X	An optional data matrix of numeric covariates, used to estimate latent regression models.
group	A vector indicating the group membership of each individual. The groups should be coded with integers starting from 0. The group denoted by 0 will be considered the reference group and the mean and variance of the latent variables in this group will by default not be estimated.
parequal	A list of seven lists indicating equality restrictions of model parameters. Entry 1 concerns restrictions on the slope parameters, entry 2 restrictions on the intercept parameters, entry 3 restrictions on scale parameters, entry 4 restrictions on regression parameters, entry 5 restrictions on mean parameters, entry 6 restrictions on variance parameters, and entry 7 restrictions on covariance parameters. Restrictions on slope parameters requires specifying which variables and dimension the restriction applies to, while restrictions on intercept and scale parameters only requires specifying which variables the restriction applies to. Restrictions on regression parameters is not yet implemented. Restrictions on variances and covariances requires specifying which covariances should be set equal, where for the variances the order is from the first to the last latent variable as defined by the model specification and where for the covariances the order is determined by the upper triangular part of the covariance matrix of the latent variables (by row). The default is to have no equality restrictions for any parameter, except those set to the same constants.
parfix	A list of lists indicating restrictions of parameters to constants. Entry 1 concerns restrictions on the slope parameters, entry 2 restrictions on the intercept parameters, entry 3 restrictions on scale parameters, entry 4 restrictions on regression parameters, entry 5 restrictions on mean parameters, entry 6 restrictions on variance parameters, and entry 7 restrictions on covariance parameters. Restrictions on slope parameters requires specifying which variables and dimension the restriction applies to and the parameter values, while restrictions on intercept and scale parameters only requires specifying which variables the restriction applies to and the parameter values. Restrictions on regression parameters is not yet implemented. Restrictions on variances and covariances requires specifying which variance and covariances should restricted and to which parameter values, where for the variances the order is from the first to the last latent variable as defined by the model specification and where for the covariances the order is determined by the upper triangular part of the covariance matrix of the latent variables (by row). The default is to have no restrictions on the slope, intercept, scale, regression, and covariance parameters while the means are set to zero and the variances are set to one. See examples for details.
groupequal	A list of lists indicating equality restrictions of parameters across groups. Entry 1 concerns restrictions on the slope parameters, entry 2 restrictions on the intercept parameters, entry 3 restrictions on scale parameters, entry 4 restrictions on regression parameters, entry 5 restrictions on mean parameters, entry 6 restrictions on variance parameters, and entry 7 restrictions on covariance parameters. Restrictions on slope parameters requires specifying which variables,

	dimension and groups the restriction applies to, while restrictions on intercept and scale parameters only requires specifying which variables and groups the restriction applies to. Restrictions on regression parameters is not yet implemented. Restrictions on variances and covariances requires specifying which variances and covariances in which groups should be set equal, where for the variances the order is from the first to the last latent variable as defined by the model specification and where for the covariances the order is determined by the upper triangular part of the covariance matrix of the latent variables (by row). The default is to have equality restrictions for all slope, intercept, and scale parameters across all groups, and to have no restrictions on the means, variances and covariance parameters across groups. See examples for details.
sw	A numeric vector of survey weights, with length equal to the number of observations.
method	A string to determine which approximation is used. "lap" - Laplace approximation. "ghq" - Gauss-Hermite quadrature. With more than four latent variables "ghq" becomes too computationally demanding to use and with more than 20 latent variables, second-order Laplace becomes too computationally demanding to use.
accuracy	An integer indicating the accuracy of the approximations used. If the Laplace approximation is used, it indicates the order of the Laplace approximation (first-order (1) and second-order (2) Laplace approximations are supported). If Gauss-Hermite quadrature is used, it indicates the number of quadrature points per latent dimension.
tol	The absolute tolerance between successive iterations to use when estimating the model parameters. Default is 1e-4.
maxit	The maximum number of iterations allowed for the parameter estimation. Default is 500.
optimizer	A character vector denoting which optimization routine to use. Options are "BFGS" (default; quasi-Newton with the BFGS Hessian approximation) and "BHHH" (quasi-Newton with the empirical cross-product Hessian approximation). "BHHH" is the fastest option but has problems converging with small sample sizes (n < 1000). "BFGS" is more robust but slower than "BHHH" and may need multiple restarts with different starting values to converge to a local maximum.
obsinfo	A logical argument indicating if the observed information matrix should be calculated after convergence, using numerical differentiation of the approximated observed gradient. (Warning: the calculation is computationally intensive with many model parameters, especially with modeupdate = TRUE.)
resdim	An optional numerical vector indicating which latent variables are to be considered residual variables, for which the covariance parameters should not be estimated.
adapt	A logical argument, indicating whether adaptive Gauss-Hermite quadrature (TRUE; default) or ordinary Gauss-Hermite quadrature (FALSE) is used.
fullexp	A logical argument, indicating whether fully exponential approximation (TRUE; default) or ordinary approximation (FALSE) is used. If method = "ghq", adapt =

FALSE, and fullexp = FALSE, both the numerator and denominator in the gradient is approximated by ordinary Gauss-Hermite quadrature. If method = "ghq", adapt = TRUE, and fullexp = FALSE, both the numerator and denominator in the gradient is approximated by adaptive Gauss-Hermite quadrature. If method = "ghq", adapt = TRUE, and fullexp = TRUE, the gradient is approximated by fully exponential Gauss-Hermite quadrature. Further, if accuracy = 1, the fully exponential Laplace approximation is used.

z.tol   The relative tolerance to use when estimating the mode of the latent variables. Default is 1e-8.

maxdiff   The maximum allowed parameter estimate update in each step of the algorithm. To stabilize the algorithm at the initial stage, this argument will limit how much each parameter estimate is allowed to change across successive iterations. Default is 0.25.

step.length   A numeric vector of length two, indicating the step lengths used. The first value is used for the first.step number of iterations and the second value for the remaining iterations. The defaults are 0.5 and 1, respectively.

first.step   The number of iterations which use the step length indicated by the first value of step.length. Default is 25, which means that the first 25 iterations will use the step length equal to the first vector entry in argument step.length and thereafter the second entry in the argument step.length will be used.

inithess   Sets the initial approximation to the Hessian matrix in the BFGS algorithm. Options are "crossprod" (empirical cross-product matrix), "identity" (identity matrix) and "NCW" (p. 142 in Nocedal and Wright (2006)). Default is "NCW". Option "crossprod" is recommended only when informative starting values are available.

startval   An optional vector of starting values for the parameters of the model. Must be equal to the number of uniquely estimated parameters of the model. For NA entries, the default starting values for the corresponding parameters will be used.

startmode   An optional matrix of starting values for the latent variable mode estimation for each individual.

starteval   Option to evaluate the log-likelihood and gradient at the specified starting values. Overrides all other options. Default is FALSE.

checkgrad   A logical argument indicating if the analytical gradient of the approximated observed log-likelihood function should be compared to the numerical gradient (used for testing only). Default it FALSE.

modeupdate   A logical argument indicating if the mode should be estimated when calculating the numerical approximation of the observed information matrix. Default it TRUE.

verbose   A logical argument to specify if detailed diagnostic information should be printed in the console during model initialization and as the algorithm runs. Default is FALSE.

filtering   A character vector indicating the type of filtering that should be used for the second-order Laplace approximation. Currently only the default option "advanced" is available, which uses an algorithm that calculates only the unique higher-order derivatives. The argument has no effect with method = "ghq".

Value

An S4 object of class 'lamleout' that includes the following slots:

`Data`	A list containing the supplied data (y, group) and information about the data (N, m).
`Estimates`	A list with parameter estimates (par, partable), standard errors (separ), mode estimates (map), covariance/information matrices (acov, Amat, Bmat), and various other quantities and objects derived from the parameter estimates (loadings, covmat, mu, modelpars, modelpartsoptC).
`Model`	A list containing model definitions (model, covstruct, modeltype), constraints (groupequal, parequal, parfix), estimation settings (method, accuracy) and filters.
`Optim`	A list containing the log-likelihood (loglik, logliki), gradient (grad, gradi), AIC, BIC, parameter estimate history (partrace), gradient history (glltrace), number of iterations (iter), and function call (call).
`Timing`	A data frame containing timing information of data preparation (prep), estimation (est) and post-estimation computations (se).

Author(s)

Björn Andersson <bjoern.h.andersson@gmail.com> and Shaobo Jin <shaobo.jin@statistik.uu.se>

References

Andersson, B., Jin, S., and Zhang, M. (2023). Fast estimation of multiple group generalized linear latent variable models for categorical observed variables. *Computational Statistics and Data Analysis*, **182**, 1-12. <doi:10.1016/j.csda.2023.107710>

Andersson, B., and Xin, T. (2021). Estimation of latent regression item response theory models using a second-order Laplace approximation. *Journal of Educational and Behavioral Statistics*, **46(2)**, 244-265. <doi:10.3102/1076998620945199>

Huber, P., Ronchetti, E., and Victoria-Feser, M.P. (2004). Estimation of generalized linear latent variable models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, **66(4)**, 893-908. <doi:10.1111/j.1467-9868.2004.05627.x>

Jin, S., Noh, M., and Lee, Y. (2018). H-Likelihood Approach to Factor Analysis for Ordinal Data. *Structural Equation Modeling: A Multidisciplinary Journal*, **25(4)**, 530-540. <doi:10.1080/10705511.2017.1403287>

Shun, Z., and McCullagh, P. (1995). Laplace approximation of high dimensional integrals. *Journal of the Royal Statistical Society: Series B (Methodological)*, **57(4)**, 749-760. <doi:10.1111/j.2517-6161.1995.tb02060.x>

Examples

```
##### Load required package.
library(mvtnorm)

##### Data generation: binary, ordinal, count, and continuous.
#####
##### Binary: GRM
```

```
##### Ordinal: GRM
##### Count: Negative-binomial
##### Continuous: Normal
##### Four indicators of each type
#####
##### Correlated latent variables and independent-clusters model
##### Model matrix for correlated latent variables
modIND <- matrix(NA, 16, 4)
modIND[1:4,1] <- rep(1, 4)
modIND[5:8,2] <- rep(1, 4)
modIND[9:12,3] <- rep(1, 4)
modIND[13:16,4] <- rep(1, 4)

##### Residual correlations between indicators
##### Model matrix for residual correlations
modRES <- matrix(NA, 16, 4)
modRES[seq(1, 16, by = 4),1] <- 1
modRES[seq(2, 16, by = 4),2] <- 1
modRES[seq(3, 16, by = 4),3] <- 1
modRES[seq(4, 16, by = 4),4] <- 1

##### Full model matrix
modFULL <- cbind(modIND, modRES)

##### Loading matrix (slope parameters)
set.seed(1234)
aFULL <- matrix(0, 16, 8)
aFULL[1:16, 1:4][!is.na(modFULL[1:16, 1:4])] <-
    runif(sum(!is.na(modFULL[1:16, 1:4])), 1.2, 2)
aFULL[1:16, 5:8][!is.na(modFULL[1:16, 5:8])] <-
    runif(sum(!is.na(modFULL[1:16, 5:8])), 0.4, 0.6)
aFULL[9:12,3] <- aFULL[9:12,3] / 2
aFULL[13:16,4] <- aFULL[13:16,4] / 2

##### Intercepts and scale parameters
bBin <- bOrd <- bCount <- bCont <- vector("list", 4)
set.seed(123456)
for(j in 1:4){
    bBin[[j]] <- rnorm(1, 0, 0.8)
    bOrd[[j]] <- c(runif(1, 0.25, 2), runif(1, -2, -0.25))
    bCount[[j]] <- c(runif(1, 0.5, 1), runif(1, 0.25, 0.5))
    bCont[[j]] <- c(runif(1, 0.5, 1), runif(1, 0.25, 0.5))
}

blist <- c(bBin, bOrd, bCount, bCont)

### Covariance matrix
set.seed(1234567)
covIND <- matrix(1, 4, 4)
covIND[lower.tri(covIND)] <- runif(6, 0.4, 0.8)
covIND[1,] <- covIND[,1]
covIND[2,] <- covIND[,2]
covIND[3,] <- covIND[,3]
```

```
covIND[4,] <- covIND[,4]

covFULL <- diag(8)
covFULL[1:4, 1:4] <- covIND

##### latent variable generation
n <- 500
set.seed(12345678)
myz <- rmvnorm(n, sigma = covFULL)

##### Data generation
set.seed(123456789)
mydata <- DGP(a = aFULL, b = blist, modeltype = c(rep("GRM", 4),
              rep("GRM", 4), rep("negbin", 4), rep("normal",
              4)), z = myz)

##### Four-dimensional independent-clusters models.
##### We specify GRM, GPCM, Negative-binomial and Normal.
##### We estimate the model with Laplace-1, Laplace-2, and
##### adaptive quadrature with 3 quadrature points (AGHQ3).
est4DGPCMNegbinLap1 <- lamle(mydata[1:250,1:16], model = modIND,
                             modeltype = c(rep("GRM", 4),
                             rep("GPCM", 4), rep("negbin", 4),
                             rep("normal", 4)), optimizer =
                             "BFGS", accuracy = 1)
est4DGPCMNegbinLap2 <- lamle(mydata[1:250,1:16], model = modIND,
                             modeltype = c(rep("GRM", 4),
                             rep("GPCM", 4), rep("negbin", 4),
                             rep("normal", 4)), optimizer =
                             "BFGS", accuracy = 2)


est4DGPCMNegbinAGHQ3 <- lamle(mydata[,1:16], model = modIND,
                              modeltype = c(rep("GRM", 4),
                              rep("GPCM", 4), rep("negbin", 4),
                              rep("normal", 4)), optimizer =
                              "BFGS", method = "ghq", accuracy =
                              3)

##### Four-dimensional multiple group model, arbitrary groups.
##### Defaults are equality constraints across groups for all
##### slope, intercept, and scale parameters, while freely
##### estimating means, variances, and covariances.
est4DGRMNegbinMGLap2 <- lamle(mydata[,1:16], model = modIND,
                              modeltype = c(rep("GRM", 4),
                              rep("GRM", 4), rep("negbin", 4),
                              rep("normal", 4)), optimizer =
                              "BFGS", accuracy = 2, group =
                              c(rep(0, n / 2), rep(1, n / 2)))

##### Eight-dimensional residual correlation models
##### We use GRM, GPCM, Negative-binomial and Normal, and
##### Laplace-2 as the estimator
```

```
parfix8D <- vector("list", 7)
for(j in 1:7) parfix8D[[j]] <- vector("list", 1)
parfix8D[[5]][[1]] <- list(
    mean = 1:8,
    value = rep(0, 8),
    groups = 1)
parfix8D[[6]][[1]] <- list(
    variance = 1:8,
    value = rep(1, 8),
    groups = 1)
parfix8D[[7]][[1]] <- list(
    covariance = c(4:7, 10:13, 15:28),
    value = rep(0, 22),
    groups = 1)

est8DGPCMNegbinLap2 <- lamle(mydata[,1:16], model = modFULL,
                             modeltype = c(rep("GRM", 4),
                             rep("GPCM", 4), rep("negbin", 4),
                             rep("normal", 4)), parfix =
                             parfix8D, optimizer = "BFGS",
                             accuracy = 2)
```

---

lamle.compute   *Compute Output from an Estimated Latent Variable Model*

Description

A function to compute different quantities for latent variable models estimated with lamle. Currently supports computing .

Usage

```
lamle.compute(x, compute = "prob", ...)
```

Arguments

x   An estimated model from function lamle.

compute   A character vector indicating what to compute. Options are 'prob' for observed
               variable category probabilities, 'expectedscore' for observed variable expected
               score functions, 'TCC' for the test characteristic curve (i.e. the expected score
               for the sum of all observed variables), 'info' for observed variable information
               functions and 'testinfo' for test information functions (i.e. the information func-
               tion for the sum of all observed variables).

... Arguments to pass to the internal functions. Argument 'variables' specifies which variables to compute things for, argument 'group' indicates which group to compute things for and argument 'z' specifies the latent variable values to use in the computations.

Value

A numeric vector with fit statistics.

Author(s)

Björn Andersson <bjoern.h.andersson@gmail.com> and Shaobo Jin <shaobo.jin@statistik.uu.se>.

References

Andersson, B., and Xin, T. (2021). Estimation of latent regression item response theory models using a second-order Laplace approximation. *Journal of Educational and Behavioral Statistics*, **46(2)** 244-265. <doi:10.3102/107699862094519>

Jin, S., Noh, M., and Lee, Y. (2018). H-Likelihood Approach to Factor Analysis for Ordinal Data. *Structural Equation Modeling: A Multidisciplinary Journal*, **25(4)**, 530-540. <doi:10.1080/10705511.2017.1403287>

Shun, Z., and McCullagh, P. (1995). Laplace approximation of high dimensional integrals. *Journal of the Royal Statistical Society: Series B (Methodological)*, **57(4)**, 749-760. <doi:10.1111/j.2517-6161.1995.tb02060.x>

Examples

```
##### Load required package.
library(mvtnorm)

#####  Generate ordinal data from the GPCM
#####
##### Item parameter generation
set.seed(123)
GPCMa <- runif(10, 0.8, 2)
GPCMb <- vector("list", 10)
for(i in 1:10) GPCMb[[i]] <- -c(runif(1, -3, -2),
                                runif(1, -1.5, -0.5),
                                runif(1, 0, 1),
                                runif(1, 1.5, 2.5))
GMCMa2d <- matrix(0, nrow = 10, ncol = 2)
GMCMa2d[1:5, 1] <- GPCMa[1:5]
GMCMa2d[6:10, 2] <- GPCMa[6:10]

##### Latent variables in two groups
n <- 200
set.seed(1234)
covmat2d <- diag(rep(1, 2))
covmat2d[!diag(2)] <- 0.6
latmat2dA <- rmvnorm(n, c(0, 0), covmat2d)
latmat2dB <- rmvnorm(n, c(1, 1), covmat2d)
```

```
##### Observed data
set.seed(12345)
dataGPCM2d <- matrix(NA, nrow = 2 * n, ncol = 10)
dataGPCM2d[1:n, ] <- DGP(GMCMa2d, GPCMb, rep("GPCM", 10),
                          latmat2dA)
dataGPCM2d[(n + 1):(2 * n), ] <- DGP(GMCMa2d, GPCMb,
                                      rep("GPCM", 10), latmat2dB)

#####  Setup two-dimensional independent-clusters model
mydim2d <- matrix(NA, nrow = 10, ncol = 2)
mydim2d[1:5, 1] <- 1
mydim2d[6:10, 2] <- 1

#####  Lap(2)
estLap2 <- lamle(y = dataGPCM2d[1:n,], model = mydim2d,
                 modeltype = rep(c("GPCM", "GRM"), 5),
                 method = "lap", accuracy = 2, optimizer =
                 "BFGS", inithess = "crossprod")

probsDim1Lap2 <- lamle.compute(estLap2, compute = "prob", variables = 1:5)
```

---

lamle.fit   *Model Fit Statistics for an Estimated Latent Variable Model*

Description

A function to compute model fit statistics for latent variable models estimated with lamle. Currently
only computes the standardized root mean square residuals (SRMSR) from the model-implied cor-
relation matrix based on the estimated model parameters and the sample correlation matrix.

Usage

```
lamle.fit(obj, obs, N = NULL, z = NULL, seed = NULL,
          use = "complete", residmat = FALSE)
```

Arguments

obj	An estimated model from function lamle.
obs	The observed data matrix, with missing values coded as NA. Ordinal data must be coded as successive integers starting from 1.
N	A numeric scalar indicating the multiplier of the sample size in the data to use when computing the model fit statistics.
z	An optional matrix of latent variable values to use when computing the model fit statistics.
seed	The seed to use when simulating data.
use	A character vector indicating how the correlaion matrix used for the SRMSR statistic should be estimated. See function cor() for details.
residmat	Optional argument to return the full residual correlation matrix.

Value

A numeric vector with fit statistics.

Author(s)

Björn Andersson <bjoern.h.andersson@gmail.com> and Shaobo Jin <shaobo.jin@statistik.uu.se>

References

Andersson, B., and Xin, T. (2021). Estimation of latent regression item response theory models using a second-order Laplace approximation. *Journal of Educational and Behavioral Statistics*, **46(2)** 244-265. <doi:10.3102/107699862094519>

Jin, S., Noh, M., and Lee, Y. (2018). H-Likelihood Approach to Factor Analysis for Ordinal Data. *Structural Equation Modeling: A Multidisciplinary Journal*, **25(4)**, 530-540. <doi:10.1080/10705511.2017.1403287>

Shun, Z., and McCullagh, P. (1995). Laplace approximation of high dimensional integrals. *Journal of the Royal Statistical Society: Series B (Methodological)*, **57(4)**, 749-760. <doi:10.1111/j.2517-6161.1995.tb02060.x>

Examples

```
##### Load required package.
library(mvtnorm)

#####  Generate ordinal data from the GPCM
#####
##### Item parameter generation
set.seed(123)
GPCMa <- runif(10, 0.8, 2)
GPCMb <- vector("list", 10)
for(i in 1:10) GPCMb[[i]] <- -c(runif(1, -3, -2),
                                runif(1, -1.5, -0.5),
                                runif(1, 0, 1),
                                runif(1, 1.5, 2.5))
GMCMa2d <- matrix(0, nrow = 10, ncol = 2)
GMCMa2d[1:5, 1] <- GPCMa[1:5]
GMCMa2d[6:10, 2] <- GPCMa[6:10]

##### Latent variables in two groups
n <- 200
set.seed(1234)
covmat2d <- diag(rep(1, 2))
covmat2d[!diag(2)] <- 0.6
latmat2dA <- rmvnorm(n, c(0, 0), covmat2d)
latmat2dB <- rmvnorm(n, c(1, 1), covmat2d)

##### Observed data
set.seed(12345)
dataGPCM2d <- matrix(NA, nrow = 2 * n, ncol = 10)
dataGPCM2d[1:n, ] <- DGP(GMCMa2d, GPCMb, rep("GPCM", 10),
                         latmat2dA)
```

```
dataGPCM2d[(n + 1):(2 * n), ] <- DGP(GMCMa2d, GPCMb,
                                     rep("GPCM", 10), latmat2dB)

#####  Setup two-dimensional independent-clusters model
mydim2d <- matrix(NA, nrow = 10, ncol = 2)
mydim2d[1:5, 1] <- 1
mydim2d[6:10, 2] <- 1

#####  Lap(2)
estLap2 <- lamle(y = dataGPCM2d[1:n,], model = mydim2d,
                 modeltype = rep(c("GPCM", "GRM"), 5),
                 method = "lap", accuracy = 2, optimizer =
                 "BFGS", inithess = "crossprod")

lamle.fit(estLap2, dataGPCM2d[1:n,], N = 10)
```

---

lamle.plot   *Plot Output from an Estimated Latent Variable Model*

Description

A function to plot different quantities for latent variable models estimated with lamle. Currently supports plotting of category probability functions, expected score functions, observed variable information functions and test information functions.

Usage

```
lamle.plot(x, toplot = "prob", palette = "Dark 3", ...)
```

Arguments

x	An estimated model from function lamle.
toplot	A character vector indicating what to plot. Options are 'prob' for observed variable category probabilities, 'expectedscore' for observed variable expected score functions, 'TCC' for the test characteristic curve (i.e. the expected score for the sum of all observed variables), 'info' for observed variable information functions and 'testinfo' for test information functions (i.e. the information function for the sum of all observed variables).
palette	Which colour palette to use for the plot.
...	Arguments to pass to the internal functions. Argument 'variables' specifies which variables to compute things for, argument 'group' indicates which group to compute things for and argument 'z' specifies the latent variable values to use in the computations.

Value

A plot of the requested output.

Author(s)

Björn Andersson <bjoern.h.andersson@gmail.com> and Shaobo Jin <shaobo.jin@statistik.uu.se>

References

Andersson, B., and Xin, T. (2021). Estimation of latent regression item response theory models using a second-order Laplace approximation. *Journal of Educational and Behavioral Statistics*, **46(2)** 244-265. <doi:10.3102/107699862094519>

Jin, S., Noh, M., and Lee, Y. (2018). H-Likelihood Approach to Factor Analysis for Ordinal Data. *Structural Equation Modeling: A Multidisciplinary Journal*, **25(4)**, 530-540. <doi:10.1080/10705511.2017.1403287>

Shun, Z., and McCullagh, P. (1995). Laplace approximation of high dimensional integrals. *Journal of the Royal Statistical Society: Series B (Methodological)*, **57(4)**, 749-760. <doi:10.1111/j.2517-6161.1995.tb02060.x>

Examples

```
##### Load required package.
library(mvtnorm)

#####  Generate ordinal data from the GPCM
#####
##### Item parameter generation
set.seed(123)
GPCMa <- runif(10, 0.8, 2)
GPCMb <- vector("list", 10)
for(i in 1:10) GPCMb[[i]] <- -c(runif(1, -3, -2),
                               runif(1, -1.5, -0.5),
                               runif(1, 0, 1),
                               runif(1, 1.5, 2.5))
GMCMa2d <- matrix(0, nrow = 10, ncol = 2)
GMCMa2d[1:5, 1] <- GPCMa[1:5]
GMCMa2d[6:10, 2] <- GPCMa[6:10]

##### Latent variables in two groups
n <- 200
set.seed(1234)
covmat2d <- diag(rep(1, 2))
covmat2d[!diag(2)] <- 0.6
latmat2dA <- rmvnorm(n, c(0, 0), covmat2d)
latmat2dB <- rmvnorm(n, c(1, 1), covmat2d)

##### Observed data
set.seed(12345)
dataGPCM2d <- matrix(NA, nrow = 2 * n, ncol = 10)
dataGPCM2d[1:n, ] <- DGP(GMCMa2d, GPCMb, rep("GPCM", 10),
                         latmat2dA)
dataGPCM2d[(n + 1):(2 * n), ] <- DGP(GMCMa2d, GPCMb,
                                     rep("GPCM", 10), latmat2dB)

#####  Setup two-dimensional independent-clusters model
mydim2d <- matrix(NA, nrow = 10, ncol = 2)
```

```
mydim2d[1:5, 1] <- 1
mydim2d[6:10, 2] <- 1

##### Lap(2)
estLap2 <- lamle(y = dataGPCM2d[1:n,], model = mydim2d,
                 modeltype = rep(c("GPCM", "GRM"), 5),
                 method = "lap", accuracy = 2, optimizer =
                 "BFGS", inithess = "crossprod")

lamle.plot(estLap2, toplot = "TIF", variables = 1:5)
lamle.plot(estLap2, toplot = "prob", variables = 5)
```

---

lamle.predict   *Compute Latent Variable Estimates from an Estimated Latent Variable*
                               Model

Description

A function to plot different quantities for latent variable models estimated with lamle. Currently
supports plotting of category probability functions, expected score functions, observed variable
information functions and test information functions.

Usage

```
lamle.predict(x, obs, estimator = "MAP", information = NULL, z.tol = 1e-8, group = NULL)
```

Arguments

x	An estimated model from function lamle.
obs	A matrix of observations, with rows indicating cases and columns indicating variables. The number of columns and order of the columns must match the data used to estimate the model provided. If only partial data are available, specify NA for the missing data.
estimator	The estimator used for the scoring, with options "MAP" for the maximum aposteriori estimator (also known as Bayesian modal estimation or the posterior mode) and "MLE" for the maximum likelihood estimator. Note that the MLE is not defined if the responses are all in the lowest or highest categories for ordinal data. Default is "MAP".
information	Option to compute the information matrices evaluated at the latent variable estimates. Options are 'expected' for the expected information (Fisher information) and 'observed' for the observed information. Default is NULL, which avoids computing the information matrices.
z.tol	The tolerance used in the numerical optimization algorithm. Default is 1e-8.
group	A vector with length equal to the number of rows in argument 'obs' which indicates group membership of each case. Default is NULL, which sets all group memberships to the first group.

Value

A list with latent variable estimates (a matrix) in the first entry and the corresponding information matrices (a list of matrices) in the second entry.

Author(s)

Björn Andersson <bjoern.h.andersson@gmail.com> and Shaobo Jin <shaobo.jin@statistik.uu.se>

References

Andersson, B., and Xin, T. (2021). Estimation of latent regression item response theory models using a second-order Laplace approximation. *Journal of Educational and Behavioral Statistics*, **46(2)** 244-265. <doi:10.3102/107699862094519>

Jin, S., Noh, M., and Lee, Y. (2018). H-Likelihood Approach to Factor Analysis for Ordinal Data. *Structural Equation Modeling: A Multidisciplinary Journal*, **25(4)**, 530-540. <doi:10.1080/10705511.2017.1403287>

Shun, Z., and McCullagh, P. (1995). Laplace approximation of high dimensional integrals. *Journal of the Royal Statistical Society: Series B (Methodological)*, **57(4)**, 749-760. <doi:10.1111/j.2517-6161.1995.tb02060.x>

Examples

```
##### Load required package.
library(mvtnorm)

#####  Generate ordinal data from the GPCM
#####
##### Item parameter generation
set.seed(123)
GPCMa <- runif(10, 0.8, 2)
GPCMb <- vector("list", 10)
for(i in 1:10) GPCMb[[i]] <- -c(runif(1, -3, -2),
                               runif(1, -1.5, -0.5),
                               runif(1, 0, 1),
                               runif(1, 1.5, 2.5))
GMCMa2d <- matrix(0, nrow = 10, ncol = 2)
GMCMa2d[1:5, 1] <- GPCMa[1:5]
GMCMa2d[6:10, 2] <- GPCMa[6:10]

##### Latent variables in two groups
n <- 200
set.seed(1234)
covmat2d <- diag(rep(1, 2))
covmat2d[!diag(2)] <- 0.6
latmat2dA <- rmvnorm(n, c(0, 0), covmat2d)
latmat2dB <- rmvnorm(n, c(1, 1), covmat2d)

##### Observed data
set.seed(12345)
dataGPCM2d <- matrix(NA, nrow = 2 * n, ncol = 10)
dataGPCM2d[1:n, ] <- DGP(GMCMa2d, GPCMb, rep("GPCM", 10),
```

```
                               latmat2dA)
dataGPCM2d[(n + 1):(2 * n), ] <- DGP(GMCMa2d, GPCMb,
                                     rep("GPCM", 10), latmat2dB)

#####  Setup two-dimensional independent-clusters model
mydim2d <- matrix(NA, nrow = 10, ncol = 2)
mydim2d[1:5, 1] <- 1
mydim2d[6:10, 2] <- 1

#####  Lap(2)
estLap2 <- lamle(y = dataGPCM2d[1:n,], model = mydim2d,
                 modeltype = rep(c("GPCM", "GRM"), 5),
                 method = "lap", accuracy = 2, optimizer =
                 "BFGS", inithess = "crossprod")

lamle.predict(estLap2, dataGPCM2d[1:n,], estimator = "MAP")
```

---

lamle.sim   *Generate Simulated Data from an Estimated Latent Variable Model*

Description

A function to simulate observed data from a latent variable model estimated with lamle.

Usage

```
lamle.sim(obj, seed = NULL, N = 1, z = NULL)
```

Arguments

obj	An estimated latent variable model from function lamle.
seed	The seed to use when simulating data.
N	A numeric scalar indicating the multiplier of the sample size in the data to use.
z	An optional matrix of latent variable values to use when simulating the observed data.

Value

A matrix with simulated observed data.

Author(s)

Björn Andersson <bjoern.h.andersson@gmail.com> and Shaobo Jin <shaobo.jin@statistik.uu.se>

References

Andersson, B., and Xin, T. (2021). Estimation of latent regression item response theory models using a second-order Laplace approximation. *Journal of Educational and Behavioral Statistics*, **46(2)** 244-265. <doi:10.3102/107699862094519>

Jin, S., Noh, M., and Lee, Y. (2018). H-Likelihood Approach to Factor Analysis for Ordinal Data. *Structural Equation Modeling: A Multidisciplinary Journal*, **25(4)**, 530-540. <doi:10.1080/10705511.2017.1403287>

Shun, Z., and McCullagh, P. (1995). Laplace approximation of high dimensional integrals. *Journal of the Royal Statistical Society: Series B (Methodological)*, **57(4)**, 749-760. <doi:10.1111/j.2517-6161.1995.tb02060.x>

Examples

```
##### Load required package.
library(mvtnorm)

#####  Generate ordinal data from the GPCM
#####
##### Item parameter generation
set.seed(123)
GPCMa <- runif(10, 0.8, 2)
GPCMb <- vector("list", 10)
for(i in 1:10) GPCMb[[i]] <- -c(runif(1, -3, -2),
                                runif(1, -1.5, -0.5),
                                runif(1, 0, 1),
                                runif(1, 1.5, 2.5))
GMCMa2d <- matrix(0, nrow = 10, ncol = 2)
GMCMa2d[1:5, 1] <- GPCMa[1:5]
GMCMa2d[6:10, 2] <- GPCMa[6:10]

##### Latent variables in two groups
n <- 200
set.seed(1234)
covmat2d <- diag(rep(1, 2))
covmat2d[!diag(2)] <- 0.6
latmat2dA <- rmvnorm(n, c(0, 0), covmat2d)
latmat2dB <- rmvnorm(n, c(1, 1), covmat2d)

##### Observed data
set.seed(12345)
dataGPCM2d <- matrix(NA, nrow = 2 * n, ncol = 10)
dataGPCM2d[1:n, ] <- DGP(GMCMa2d, GPCMb, rep("GPCM", 10),
                         latmat2dA)
dataGPCM2d[(n + 1):(2 * n), ] <- DGP(GMCMa2d, GPCMb,
                                     rep("GPCM", 10), latmat2dB)

#####  Setup two-dimensional independent-clusters model
mydim2d <- matrix(NA, nrow = 10, ncol = 2)
mydim2d[1:5, 1] <- 1
mydim2d[6:10, 2] <- 1

#####  Lap(2)
```

```
estLap2 <- lamle(y = dataGPCM2d[1:n,], model = mydim2d,
                 modeltype = rep(c("GPCM", "GRM"), 5),
                 method = "lap", accuracy = 2, optimizer =
                 "BFGS", inithess = "crossprod")

lamle.sim(estLap2, N = 1)
```

---

lamleout-class   *Class "lamleout"*

Description

Output for an estimated generalized linear latent variable model.

Objects from the Class

Objects can be created by calls of the form new("lamleout", ...).

Slots

Data A list containing the supplied data (y, group) and information about the data (N, m).

Estimates A list with parameter estimates (par, partable), standard errors (separ), mode estimates (map), covariance/information matrices (acov, Amat, Bmat), and various other quantities and objects derived from the parameter estimates (loadings, covmat, mu, modelpars, modelpart-soptC).

Model A list containing model definitions (model, covstruct, modeltype), constraints (groupequal, parequal, parfix), estimation settings (method, accuracy) and filters.

Optim A list containing the log-likelihood (loglik, logliki), gradient (grad, gradi), AIC, BIC, parameter estimate history (partrace), gradient history (glltrace), number of iterations (iter), and function call (call).

Timing A data frame containing timing information of data preparation (prep), estimation (est) and post-estimation computations (se).

Author(s)

<bjorn.andersson@cemo.uio.no>

See Also

[lamle](#)

Examples

```
showClass("lamleout")
```

Index