

Package ‘LatentBMA’

June 27, 2024

Type Package

Title Bayesian Model Averaging for Univariate Link Latent Gaussian Models

Version 0.1.0

Author Gregor Zens [aut, cre],
Mark F.J. Steel [aut]

Maintainer Gregor Zens <zens@iiasa.ac.at>

Description Bayesian model averaging (BMA) algorithms for univariate link latent Gaussian models (ULLGMs). For detailed information, refer to Steel M.F.J. & Zens G. (2024) “Model Uncertainty in Latent Gaussian Models with Univariate Link Function” <doi:10.48550/arXiv.2406.17318>. The package supports various g-priors and a beta-binomial prior on the model space. It also includes auxiliary functions for visualizing and tabulating BMA results. Currently, it offers an out-of-the-box solution for model averaging of Poisson log-normal (PLN) and binomial logistic-normal (BiL) models. The codebase is designed to be easily extendable to other likelihoods, priors, and link functions.

Encoding UTF-8

License MIT + file LICENSE

RoxygenNote 7.3.1

Imports ggplot2 (>= 3.5.1), knitr (>= 1.47), mnormt (>= 2.1.1),
progress (>= 1.2.3), reshape2 (>= 1.4.4)

Suggests rmarkdown

VignetteBuilder knitr

NeedsCompilation no

Repository CRAN

Date/Publication 2024-06-27 14:40:26 UTC

Contents

plotBeta	2
plotModelSize	3
plotPIP	4

summaryBMA	5
topModels	6
tracePlot	7
ULLGM_BMA	7

Index	10
--------------	-----------

plotBeta	<i>Visualization of Posterior Means of Coefficients</i>
----------	---

Description

plotBeta produces a visualization of the estimated posterior means of the coefficients, extracted from ULLGM_BMA results.

Usage

```
plotBeta(x,
         variable_names = NULL,
         sort           = TRUE)
```

Arguments

x	The output object of ULLGM_BMA.
variable_names	A character vector specifying the names of the columns of X.
sort	Logical, indicating whether the plot should be sorted by posterior mean. Defaults to TRUE.

Value

Returns a 'ggplot2::ggplot' object.

Author(s)

Gregor Zens

Examples

```
# Load package
library(LatentBMA)

# Example: Estimate a PLN model under a BRIC prior with m = p/2 using simulated data
# Note: Use more samples for actual analysis
# Note: nsave = 250 and nburn = 250 are for demonstration purposes
X <- matrix(rnorm(100*20), 100, 20)
z <- 2 + X %*% c(0.5, -0.5, rep(0, 18)) + rnorm(100, 0, sqrt(0.25))
y <- rpois(100, exp(z))
results_pln <- ULLGM_BMA(X = X, y = y, model = "PLN", nsave = 250, nburn = 250)
plotBeta(results_pln)
```

plotModelSize	<i>Visualization of Model Size Posterior Distribution</i>
---------------	---

Description

plotModelSize produces a visualization of the posterior distribution of model size, extracted from ULLGM_BMA results.

Usage

```
plotModelSize(x)
```

Arguments

x The output object of ULLGM_BMA.

Value

Returns a 'ggplot2::ggplot' object visualizing the posterior distribution of model size.

Author(s)

Gregor Zens

Examples

```
# Load package
library(LatentBMA)

# Example: Estimate a PLN model under a BRIC prior with m = p/2 using simulated data
# Note: Use more samples for actual analysis
# Note: nsave = 250 and nburn = 250 are for demonstration purposes
X <- matrix(rnorm(100*20), 100, 20)
z <- 2 + X %*% c(0.5, -0.5, rep(0, 18)) + rnorm(100, 0, sqrt(0.25))
y <- rpois(100, exp(z))
results_pln <- ULLGM_BMA(X = X, y = y, model = "PLN", nsave = 250, nburn = 250)
plotModelSize(results_pln)
```

`plotPIP`*Visualization of Posterior Inclusion Probabilities*

Description

`plotPIP` produces a visualization of the posterior inclusion probabilities (PIPs) extracted from ULLGM_BMA results.

Usage

```
plotPIP(x,  
        variable_names = NULL,  
        sort            = TRUE)
```

Arguments

`x` The output object of ULLGM_BMA.
`variable_names` A character vector specifying the names of the columns of X.
`sort` Logical, indicating whether the plot should be sorted by PIP. Defaults to TRUE.

Value

Returns a 'ggplot2::ggplot' object.

Author(s)

Gregor Zens

Examples

```
# Load package  
library(LatentBMA)  
  
# Example: Estimate a PLN model under a BRIC prior with m = p/2 using simulated data  
# Note: Use more samples for actual analysis  
# Note: nsave = 250 and nburn = 250 are for demonstration purposes  
X <- matrix(rnorm(100*20), 100, 20)  
z <- 2 + X %*% c(0.5, -0.5, rep(0, 18)) + rnorm(100, 0, sqrt(0.25))  
y <- rpois(100, exp(z))  
results_pln <- ULLGM_BMA(X = X, y = y, model = "PLN", nsave = 250, nburn = 250)  
plotPIP(results_pln)
```

summaryBMA

*Summary Tables for ULLGM_BMA Estimation Results***Description**

summaryBMA produces a table with estimated posterior means, standard deviations, and posterior inclusion probabilities (PIPs) for the results of a ULLGM_BMA estimation.

Usage

```
summaryBMA(x,
           variable_names = NULL,
           digits          = 3,
           sort            = FALSE,
           type             = "pandoc")
```

Arguments

`x` The output object of ULLGM_BMA.

`variable_names` A character vector specifying the names of the columns of X.

`digits` Number of digits to round the table to. Defaults to 3.

`sort` Logical, indicating whether the table should be sorted by PIPs. Default is FALSE.

`type` A character string indicating the format of the table. Options are 'pandoc' (default), 'latex', or 'html'.

Value

Returns a 'knitr::kable' object containing the summary table.

Author(s)

Gregor Zens

Examples

```
# Load package
library(LatentBMA)

# Example: Estimate a PLN model under a BRIC prior with m = p/2 using simulated data
# Note: Use more samples for actual analysis
# Note: nsave = 250 and nburn = 250 are for demonstration purposes
X <- matrix(rnorm(100*20), 100, 20)
z <- 2 + X %*% c(0.5, -0.5, rep(0, 18)) + rnorm(100, 0, sqrt(0.25))
y <- rpois(100, exp(z))
results_pln <- ULLGM_BMA(X = X, y = y, model = "PLN", nsave = 250, nburn = 250)
summaryBMA(results_pln)
```

topModels

Extract Top Models from ULLGM_BMA Estimation Results

Description

topModels produces a table of the top n models from a ULLGM_BMA object, sorted by posterior model probabilities.

Usage

```
topModels(x,
          variable_names = NULL,
          type            = "pandoc",
          digits         = 3,
          n              = 5)
```

Arguments

x The output object of ULLGM_BMA.

variable_names A character vector specifying the names of the columns of X.

type A character string indicating the format of the table. Options are 'pandoc' (default), 'latex', or 'html'.

digits Number of digits to round the table to. Defaults to 3.

n Number of top models to be returned. Defaults to 5.

Value

Returns a 'knitr::kable' object containing the table of top models.

Author(s)

Gregor Zens

Examples

```
# Load package
library(LatentBMA)

# Example: Estimate a PLN model under a BRIC prior with m = p/2 using simulated data
# Note: Use more samples for actual analysis
# Note: nsave = 250 and nburn = 250 are for demonstration purposes
X <- matrix(rnorm(100*20), 100, 20)
z <- 2 + X %*% c(0.5, -0.5, rep(0, 18)) + rnorm(100, 0, sqrt(0.25))
y <- rpois(100, exp(z))
results_pln <- ULLGM_BMA(X = X, y = y, model = "PLN", nsave = 250, nburn = 250)
# Top 5 models
topModels(results_pln)
```

tracePlot	<i>Traceplots for Selected Parameters</i>
-----------	---

Description

tracePlot produces traceplots for selected parameters, extracted from ULLGM_BMA results.

Usage

```
tracePlot(x, parameter = "beta", index = 1)
```

Arguments

x	The output object of ULLGM_BMA.
parameter	Specifies which parameter should be considered for the traceplot. Options are "beta" for coefficients, "alpha" for the intercept (the default), "modelsize" for model size, and "sigma2" for the error variance.
index	If parameter = "beta", specifies which coefficient should be shown. Defaults to 1, corresponding to the covariate in the first column of X.

Value

Returns a 'ggplot2::ggplot' object.

Author(s)

Gregor Zens

ULLGM_BMA	<i>Bayesian Model Averaging for Poisson Log-Normal and Binomial Logistic-Normal Regression Models</i>
-----------	---

Description

ULLGM_BMA estimates Bayesian regression models using either a Poisson log-normal (PLN) or binomial logistic-normal (BiL) regression framework. It accounts for model uncertainty via Bayesian model averaging.

Usage

```
ULLGM_BMA(X,
           y,
           model = "PLN",
           gprior = "BRIC",
           nsave = 10000,
           nburn = 2000,
           Ni = NULL,
           m = NULL,
           verbose = TRUE)
```

Arguments

X	A $n \times p$ design matrix where n is the number of observations and p is the number of explanatory variables.
y	A $n \times 1$ response vector. For PLN and BiL models, this is a count response.
model	Indicates the model to be estimated. Options are "PLN" for the Poisson log-normal model and "BiL" for the binomial logistic-normal model. Default is "PLN".
gprior	Specifies the g-prior to be used. Options under fixed g are "BRIC" ($g = \max(n, p^2)$), "UIP" ($g = n$), "RIC" ($g = p^2$), "SQRT-N" ($g = \sqrt{n}$). Options under random g are the hyper-g and hyper-g/n priors of Liang et al. (2008) ("hyper-g(a=3)", "hyper-g(a=4)", "hyper-g/n(a=3)", "hyper-g/n(a=4)"), the prior of Zellner & Siow (1980) ("zellnersiow"), and a Beta(0.5, 0.5) prior on $g/(1+g)$ ("horseshoe"). Default is "BRIC".
nsave	The number of saved posterior samples. Defaults to 10,000.
nburn	The number of initial burn-in samples. Defaults to 2,000.
Ni	A vector containing the number of trials for each observation when estimating a binomial logistic-normal model. Required if model = "BiL".
m	The prior expected model size as per the beta-binomial prior of Ley and Steel (2009). Defaults to $p/2$, representing a uniform prior on model size.
verbose	Logical indicator of whether progress should be printed during estimation. Default is TRUE.

Value

A list containing the inputs and selected posterior simulation outputs, such as posterior chains for the coefficients and inclusion vectors.

Note

All explanatory variables in X are automatically demeaned within the function. All models do automatically include an intercept term.

References

- Liang, F., Paulo, R., Molina, G., Clyde, M. A., & Berger, J. O. (2008). Mixtures of g priors for Bayesian variable selection. *Journal of the American Statistical Association*, 103(481), 410-423.
- Zellner, A., & Siow, A. (1980). Posterior odds ratios for selected regression hypotheses. *Trabajos de estadística y de investigación operativa*, 31, 585-603.
- Ley, E., & Steel, M. F. J. (2009). On the effect of prior assumptions in Bayesian model averaging with applications to growth regression. *Journal of Applied Econometrics*, 24(4), 651-674.

Examples

```
# Load package
library(LatentBMA)

# Example 1: Estimate a PLN model under a BRIC prior with m = p/2 using simulated data
# Note: Use more samples for actual analysis
# Note: nsave = 250 and nburn = 250 are for demonstration purposes
X <- matrix(rnorm(100*20), 100, 20)
z <- 2 + X %*% c(0.5, -0.5, rep(0, 18)) + rnorm(100, 0, sqrt(0.25))
y <- rpois(100, exp(z))
results_pln <- ULLGM_BMA(X = X, y = y, model = "PLN", nsave = 250, nburn = 250)

# Example 2: Estimate a BiL model under a Zellner-Siow prior with m = 5 using simulated data
# Note: Use more samples for actual analysis
# Note: nsave = 250 and nburn = 250 are for demonstration purposes
X <- matrix(rnorm(100*20), 100, 20)
Ni <- rep(50, 100)
z <- 2 + X %*% c(0.5, -0.5, rep(0, 18)) + rnorm(100, 0, sqrt(0.25))
y <- rbinom(100, Ni, 1 / (1 + exp(-z)))
results_bil <- ULLGM_BMA(X = X, y = y, Ni = Ni, model = "BiL", nsave = 250, nburn = 250,
                        m = 5, gprior = "zellnersiow")
```

Index

plotBeta, [2](#)
plotModelSize, [3](#)
plotPIP, [4](#)

summaryBMA, [5](#)

topModels, [6](#)
tracePlot, [7](#)

ULLGM_BMA, [7](#)