Package ‘MBSGS’

January 20, 2025

Type Package

Title Multivariate Bayesian Sparse Group Selection with Spike and Slab

Version 1.1.0

Author Benoit Liquet, Matthew Sutton

Maintainer Benoit Liquet <benoit.liquet@qut.edu.au>

Imports MCMCpack, MASS, mgcv, mnormt, truncnorm

Description An implementation of a Bayesian sparse group model using spike and slab priors in a regression context. It is designed for regression with a multivariate response variable, but also provides an implementation for univariate response.

License GPL (>= 2.0)

Encoding UTF-8

NeedsCompilation no

Repository CRAN

Date/Publication 2017-08-09 08:06:21 UTC

Contents

BGLSS *Bayesian Group Lasso with Spike and Slab prior*

Description

Run a gibbs sampler for a Bayesian group lasso model with spike and slab prior. This function is designed for an univariate response model and when the design matrix has a group structur. Run a gibbs sampler for a Bayesian group lasso model with spike and slab prior. This function is designed for an univariate response model and when the design matrix has a group structure.

Usage

```
BGLSS(Y, X, niter = 10000, burnin = 5000, group_size, a = 1,
b = 1, num_update = 100, niter.update = 100, verbose = FALSE,
alpha = 0.1, gamma = 0.1, pi_prior = TRUE, pi = 0.5,
update_tau = TRUE, option.weight.group = FALSE,
option.update = "global", lambda2_update = NULL)
```

Arguments

`Y`	A numerical vector representing the univariate response variable.
`X`	A matrix respresenting the design matrix of the linear regression model.
`niter`	Number of iteration for the Gibbs sampler.
`burnin`	Number of burnin iteration
`group_size`	Integer vector representing the size of the groups of the design matrix `X`
`a`	First shape parameter of the conjugate beta prior for `pi_0`. Default is 1.
`b`	Second shape parameter of the conjugate beta prior for `pi_0`. Default is 1.
`num_update`	Number of update regarding the scaling of the shrinkage parameter lambda which is calibrated by a Monte Carlo EM algorithm
`niter.update`	Number of itertion regarding the scaling of the shrinkage parameter lambda which is calibrated by a Monte Carlo EM algorithm
`verbose`	Logical. If "TRUE" iterations are displayed.
`alpha`	Shape parameter of the Inverse Gamma prior on the variance of the noise for the linear regression model.
`gamma`	Scale parameter of the Inverse Gamma prior on the variance of the noise for the linear regression model.
`pi_prior`	Logical. If "TRUE" a beta prior is used for pi
`pi`	Initial value for pi_0 which will be updated if `pi_prior`="TRUE""
`update_tau`	Logical. If "TRUE" then a Monte Carlo EM algorithm is used to update lambda
`option.weight.group`	If "TRUE" then the group size is used for shrinkage penalty purpose.
`option.update`	Two options are proposed for updating lambda. A "Local" update or a "Global" update
`lambda2_update`	Value of the square of lambda when `update_tau`="FALSE"

Value

BGLSS returns a list that contains the following components:

pos_mean	The posterior mean estimate of the regression coefficients
pos_median	The posterior mean estimate of the regression coefficients
coef	A matrix with the regression coefficients sampled at each iteration

Author(s)

Benoit Liquet, Matthew Sutton and Xiaofan Xu.

References

B. Liquet, K. Mengersen, A. Pettitt and M. Sutton. (2016). Bayesian Variable Selection Regression Of Multivariate Responses For Group Data. *Submitted in Bayesian Analysis*.

Xu, X. and Ghosh, M. (2015). Bayesian Variable Selection and Estimation for Group Lasso. *Bayesian Analysis*, 10(4): 909–936.

See Also

BSGSSS

Examples

```
## Simulation of datasets X and Y with group variables
set.seed(1)
data1 = gen_data_uni(nsample = 120,cor.var=0.5, ntrain = 80)
data1 = normalize(data1)


true_model <- data1$true_model
X <- data1$X
Y<- data1$Y
train_idx <- data1$train_idx
gsize <- data1$gsize
## We recommend to set niter=50000, burnin=10000
## num_update = 100 and niter.update = 100
## to reach convergence

model <- BGLSS(Y[,1],X,niter=500,burnin=100,group_size=gsize,
num_update = 20,niter.update = 20)
model$pos_median!=0
true_model
```

BSGSSS *Bayesian Sparse Group Selection with Spike and Slab priors*

Description

Run a gibbs sampler for a Bayesian sparse group selection model with spike and slab priors. This function is designed for an univariate response model and when the design matrix has a group structure.

Usage

```
BSGSSS(Y, X, group_size, niter = 10000, burnin = 5000,
pi0 = 0.5, pi1 = 0.5, num_update = 100, niter.update = 100,
alpha = 0.1, gamma = 0.1, a1 = 1, a2 = 1, c1 = 1, c2 = 1,
pi_prior = TRUE)
```

Arguments

Y	A numerical vector representing the univariate response variable.
X	A matrix respresenting the design matrix of the linear regression model.
group_size	Integer vector representing the size of the groups of the design matrix X
niter	Number of iteration for the Gibbs sampler.
burnin	Number of burnin iteration
pi0	Initial value for pi0 which will be updated if pi_prior="TRUE"
pi1	Initial value for pi1 which will be updated if pi_prior="TRUE"
num_update	Number of update regarding the scaling of the shrinkage parameter lambda which is calibrated by a Monte Carlo EM algorithm
niter.update	Number of itertion regarding the scaling of the shrinkage parameter lambda which is calibrated by a Monte Carlo EM algorithm
alpha	Shape parameter of the Inverse Gamma prior on the variance of the noise for the linear regression model.
gamma	Scale parameter of the Inverse Gamma prior on the variance of the noise for the linear regression model.
a1	First shape parameter of the conjugate beta hyper-prior for pi_0. Default is 1.
a2	Second shape parameter of the conjugate beta prior for pi_0. Default is 1.
c1	First shape parameter of the conjugate beta hyper-prior for pi_1. Default is 1.
c2	Second shape parameter of the conjugate beta prior for pi_1. Default is 1.
pi_prior	Logical. If "TRUE" beta priors are used for pi0 and pi1

Value

BSGSSS returns a list that contains the following components:

pos_mean	The posterior mean estimate of the regression coefficients
pos_median	The posterior mean estimate of the regression coefficients
coef	A matrix with the regression coefficients sampled at each iteration

Author(s)

Benoit Liquet, Matthew Sutton and Xiaofan Xu.

References

B. Liquet, K. Mengersen, A. Pettitt and M. Sutton. (2016). Bayesian Variable Selection Regression Of Multivariate Responses For Group Data. *Submitted in Bayesian Analysis*.

Xu, X. and Ghosh, M. (2015). Bayesian Variable Selection and Estimation for Group Lasso. *Bayesian Analysis*, 10(4): 909–936.

See Also

BGLSS

Examples

```
## Simulation of datasets X and Y with group variables
set.seed(1)
data1 = gen_data_uni(nsample = 120,cor.var=0.5, ntrain = 80)
data1 = normalize(data1)

true_model <- data1$true_model
X <- data1$X
Y<- data1$Y
train_idx <- data1$train_idx
gsize <- data1$gsize
## We recommend to set niter=50000, burnin=10000
## num_update = 100 and niter.update = 100
## to reach convergence

model <- BSGSSS(Y[,1],X,niter=500,burnin=100,group_size=gsize,
num_update = 20,niter.update = 20)
model$pos_median!=0
true_model
```

MBGLSS *Multivariate Bayesian Group Lasso with Spike and Slab prior*

Description

Run a gibbs sampler for a Multivariate Bayesian group lasso model with spike and slab prior. This function is designed for a regression model with multivariate response, where the design matrix has a group structure.

Usage

```
MBGLSS(Y, X, niter = 10000, burnin = 5000, group_size,
a = 1, b = 1, num_update = 100, niter.update = 100,
verbose = FALSE, pi_prior = TRUE, pi = 0.5,
d = 3, update_tau = TRUE, option.update = "global")
```

Arguments

`Y`	A numerical vector representing the univariate response variable.
`X`	A matrix respresenting the design matrix of the linear regression model.
`niter`	Number of iteration for the Gibbs sampler.
`burnin`	Number of burnin iteration
`group_size`	Integer vector representing the size of the groups of the design matrix `X`
`a`	First shape parameter of the conjugate beta prior for `pi_0`. Default is 1.
`b`	Second shape parameter of the conjugate beta prior for `pi_0`. Default is 1.
`num_update`	Number of update regarding the scaling of the shrinkage parameter lambda which is calibrated by a Monte Carlo EM algorithm
`niter.update`	Number of itertion regarding the scaling of the shrinkage parameter lambda which is calibrated by a Monte Carlo EM algorithm
`verbose`	Logical. If "TRUE" iterations are displayed.
`pi_prior`	Logical. If "TRUE" a beta prior is used for pi
`pi`	Initial value for pi_0 which will be updated if `pi_prior`="TRUE""
`d`	Degree of freedom of the inverse Wishart prior of the covariance matrix of the response variable. By default `d` is set to 3.
`update_tau`	Logical. If "TRUE" then a Monte Carlo EM algorithm is used to update lambda
`option.update`	Two options are proposed for updating lambda. A "Local" update or a "Global" update

Value

`BSGSSS` returns a list that contains the following components:

`pos_mean`	The posterior mean estimate of the regression coefficients
`pos_median`	The posterior mean estimate of the regression coefficients
`coef`	A matrix with the regression coefficients sampled at each iteration

Author(s)

Benoit Liquet and Matthew Sutton.

References

B. Liquet, K. Mengersen, A. Pettitt and M. Sutton. (2016). Bayesian Variable Selection Regression Of Multivariate Responses For Group Data. *Submitted in Bayesian Analysis*.

See Also

MBSGSSS

Examples

```
## Not run:
## Simulation of datasets X and Y with group variables
data1 = gen_data_Multi(nsample = 120, ntrain = 80)
data1 = Mnormalize(data1)

true_model <- data1$true_model
X <- data1$X
Y<- data1$Y
train_idx <- data1$train_idx
gsize <- data1$gsize
niter <- 2000
burnin <- 1000

model <- MBGLSS(Y,X,niter,burnin,gsize,num_update = 100,
niter.update = 100)
model$pos_median[,1]!=0

## End(Not run)
```

MBSGS-internal *Internal Functions*

Description

Internal functions not to be used by the user.

MBSGSSS — *Multivariate Bayesian Sparse Group Selection with Spike and Slab priors*

Description

Run a gibbs sampler for a Multivariate Bayesian sparse group selection model with spike and slab prior. This function is designed for a regression model with multivariate response, where the design matrix has a group structure.

Usage

```
MBSGSSS(Y, X, group_size, pi0 = 0.5, pi1 = 0.5,
a1 = 1, a2 = 1, c1 = 1,c2 = 1, pi_prior = TRUE,
niter = 10000, burnin = 5000, d = 3,
num_update = 100, niter.update = 100)
```

Arguments

`Y`	A numerical vector representing the univariate response variable.
`X`	A matrix respresenting the design matrix of the linear regression model.
`group_size`	Integer vector representing the size of the groups of the design matrix `X`
`pi0`	Initial value for pi0 which will be updated if `pi_prior`="TRUE"
`pi1`	Initial value for pi1 which will be updated if `pi_prior`="TRUE"
`a1`	First shape parameter of the conjugate beta hyper-prior for `pi_0`. Default is 1.
`a2`	Second shape parameter of the conjugate beta prior for `pi_0`. Default is 1.
`c1`	First shape parameter of the conjugate beta hyper-prior for `pi_1`. Default is 1.
`c2`	Second shape parameter of the conjugate beta prior for `pi_1`. Default is 1.
`pi_prior`	Logical. If "TRUE" beta priors are used for pi0 and pi1
`niter`	Number of iteration for the Gibbs sampler.
`burnin`	Number of burnin iteration
`d`	Degree of freedom of the inverse Wishart prior of the covariance matrix of the response variable. By default `d` is set to 3.
`num_update`	Number of update regarding the scaling of the shrinkage parameter lambda which is calibrated by a Monte Carlo EM algorithm
`niter.update`	Number of itertion regarding the scaling of the shrinkage parameter lambda which is calibrated by a Monte Carlo EM algorithm

Author(s)

Benoit Liquet and Matthew Sutton.

References

B. Liquet, K. Mengersen, A. Pettitt and M. Sutton. (2016). Bayesian Variable Selection Regression Of Multivariate Responses For Group Data. *Submitted in Bayesian Analysis*.

See Also

MBGLSS

Examples

```
## Not run:
## Simulation of datasets X and Y with group variables
data1 = gen_data_Multi(nsample = 120, ntrain = 80)
data1 = Mnormalize(data1)

true_model <- data1$true_model
X <- data1$X
Y<- data1$Y
train_idx <- data1$train_idx
gsize <- data1$gsize
niter <- 2000
burnin <- 1000

model <- MBSGSSS(Y,X,niter=niter,burnin=burnin,group_size=gsize,
num_update = 50,niter.update = 50)
model$pos_median[,1]!=0

## End(Not run)
```

Index