Package 'ghorseshoe'

April 23, 2019

Type Package

Title Gibbs sampler for grouped horseshoe regression

Version 1.0

Date 2019-04-23

Author Jonathan Boss

Maintainer Jonathan Boss <bossjona@umich.edu>

Description This package implements a Gibbs sampler for grouped horsehose regression. The main function is the grouped_horseshoe function which calls the grouped_horseshoe_gibbs_sampler function implemented C++.

License GPL-3

Imports Rcpp (>= 1.0.0), GeneralizedHyperbolic

LinkingTo Rcpp, RcppArmadillo

RoxygenNote 6.1.1

R topics documented:

`ghorseshoe-package` *Gibbs sampler for grouped horseshoe regression*

Description

This package implements a Gibbs sampler for grouped horsehose regression. The main function is the grouped_horseshoe function which calls the grouped_horseshoe_gibbs_sampler function implemented C++.

Details

The DESCRIPTION file: This package was not yet installed at build time.

Index: This package was not yet installed at build time.
This package implements a Gibbs sampler for grouped horseshoe regression. There are three functions that are included in the package. The primary function is the grouped_horseshoe function which only requires four inputs: (1) the outcome vector (Y), (2) the part of the design matrix corresponding to the adjustment covariates (C), (3) the part of the design matrix corresponding to the variables we want to implement grouped horseshoe shrinkage on (X), and (4) the grp_idx argument which lists which covariate group each column of X belongs. For the function to work properly, make sure that group indicies in grp_idx take the values 1,2,...,J (where J is the total number of covariate groups in X) and that the columns of X are ordered by grp_idx (that is, all the columns in X corresponding to group 1 come first, followed by all the columns in X corresponding to group 2, and so on). The grouped_horseshoe_gibbs_sampler function is called by the grouped_horseshoe function and runs the grouped horsehoe regression Gibbs sampler. The rgig_use function calls the rgig function from the GeneralizedHyperbolic package in R which is used in the grouped_horseshoe_gibbs_sampler function. Therefore, the only function that the user needs to use is the grouped_horseshoe function.

Author(s)

Jonathan Boss

Maintainer: Jonathan Boss <bossjona@umich.edu>

`grouped_horseshoe` *Grouped horseshoe regression Gibbs sampler.*

Description

Grouped horseshoe regression Gibbs sampler.

Usage

```
grouped_horseshoe(X, C, Y, grp_idx, alpha_inits = rep(0, ncol(C)),
  beta_inits = rep(0, ncol(X)), lambda_sq_inits = rep(1, ncol(X)),
  gamma_sq_inits = rep(1, length(unique(grp_idx))), eta_inits = rep(1,
  length(unique(grp_idx))), mu_init = 0, tau_sq_init = 1,
  sigma_sq_init = 1, psi_sq_init = 1, nu_init = 1, n_burn_in = 500,
  n_samples = 1000, n_thin = 1, error_tol = 1e-07, a1 = 0.001,
  b1 = 0.001, a2 = 0.001, b2 = 0.001, verbose = TRUE)
```

Arguments

`X`	A (n x M) matrix of covariates that we want to apply grouped horseshoe shrinkage on.
`C`	A (n x K) matrix of covariates that we want to apply ridge regression shrinkage on (typically adjustment covariates).
`Y`	A (n x 1) column vector of responses.
`grp_idx`	A (1 x M) row vector indicating which group of the J groups the p covariates in X belong to.

alpha_inits	A (K x 1) column vector containing initial values for the regression coefficients corresponding to C.
beta_inits	A (M x 1) column vector containing initial values for the regression coefficients corresponding to X.
lambda_sq_inits	A (M x 1) column vector containing initial values for the local shrinkage parameters.
gamma_sq_inits	A (J x 1) column vector containing initial values for the group shrinkage parameters.
eta_inits	A (J x 1) column vector containing initial values for the mixing parameters.
mu_init	Initial value for the intercept parameter (double).
tau_sq_init	Initial value for the global shrinkage parameter (double).
sigma_sq_init	Initial value for the residual variance (double).
psi_sq_init	Initial value for the variailbity in the alphas (double).
nu_init	Initial value for the augmentation variable (double).
n_burn_in	The number of burn-in samples (integer).
n_samples	The number of posterior draws (integer).
n_thin	The thinning interval (integer).
error_tol	Parameter that controls numerical stability of the algorithm (double).
a1	Shape parameter for the inverse gamma hyperprior on psi_sq (double).
b1	Scale parameter for the inverse gamma hyperprior on psi_sq (double).
a2	Shape parameter for the gamma hyperprior on the etas (double).
b2	Scale parameter for the gamma hyperprior on the etas (double).
verbose	Boolean value which indicates whether or not to print the progress of the Gibbs sampler.

Value

A list containing the posterior draws of (1) the intercept term (mus) (2) the regression coefficients (alphas and betas) (3) the individual shrinkage parameters (lambda_sqs) (4) the group shrinkage parameters (gamma_sqs) (5) the global shrinkage parameter (tau_sqs) (6) the residual error variance (sigma_sqs) (7) the mixing parameter (etas) and (8) the variance of the prior on alpha (psi_sqs). The list also contains the specified hyperparameters (hyperparam_a1, hyperparam_b1, hyperparam_a2, and hyperparam_b2), details regarding the dataset (X, C, Y, grp_idx), and Gibbs sampler details (n_burn_in, n_samples, and n_thin).

---

grouped_horseshoe_gibbs_sampler

Gibbs sampler for the grouped horseshoe model.

Description

An Rcpp function that implements a Gibbs sampler for the grouped horseshoe model.

Usage

```
grouped_horseshoe_gibbs_sampler(X, C, Y, grp_idx, grp_size, grp_size_cs,
  alpha_inits, beta_inits, lambda_sq_inits, gamma_sq_inits, eta_inits,
  mu_init = 0, tau_sq_init = 1, sigma_sq_init = 1, psi_sq_init = 1,
  nu_init = 1, n_burn_in = 500L, n_samples = 1000L, n_thin = 1L,
  error_tol = 1e-07, a1 = 0.001, b1 = 0.001, a2 = 0.001,
  b2 = 0.001, verbose = TRUE)
```

Arguments

X	A (n x M) matrix of covariates that we want to apply grouped horseshoe shrinkage on.
C	A (n x K) matrix of covariates that we want to apply ridge regression shrinkage on (typically adjustment covariates).
Y	A (n x 1) column vector of responses.
grp_idx	A (1 x M) row vector indicating which group of the J groups the p covariates in X belong to.
grp_size	A (1 x J) row vector indicating the number of covariates in each group.
grp_size_cs	A (1 x J) row vector that is the cumulative sum of grp_size (indicating the indicies where each group ends).
alpha_inits	A (K x 1) column vector containing initial values for the regression coefficients corresponding to C.
beta_inits	A (M x 1) column vector containing initial values for the regression coefficients corresponding to X.
lambda_sq_inits	A (M x 1) column vector containing initial values for the local shrinkage parameters.
gamma_sq_inits	A (J x 1) column vector containing initial values for the group shrinkage parameters.
eta_inits	A (J x 1) column vector containing initial values for the mixing parameters.
mu_init	Initial value for the intercept parameter (double).
tau_sq_init	Initial value for the global shrinkage parameter (double).
sigma_sq_init	Initial value for the residual variance (double).
psi_sq_init	Initial value for the variailbity in the alphas (double).
nu_init	Initial value for the augmentation variable (double).
n_burn_in	The number of burn-in samples (integer).
n_samples	The number of posterior draws (integer).
n_thin	The thinning interval (integer).
error_tol	Parameter that controls numerical stability of the algorithm (double).
a1	Shape parameter for the inverse gamma hyperprior on psi_sq (double).
b1	Scale parameter for the inverse gamma hyperprior on psi_sq (double).
a2	Shape parameter for the gamma hyperprior on the etas (double).
b2	Scale parameter for the gamma hyperprior on the etas (double).
verbose	Boolean value which indicates whether or not to print the progress of the Gibbs sampler.

Value

A list containing the posterior draws of (1) the intercept term (mus) (2) the regression coefficients (alphas and betas) (3) the individual shrinkage parameters (lambda_sqs) (4) the group shrinkage parameters (gamma_sqs) (5) the global shrinkage parameter (tau_sqs) (6) the residual error variance (sigma_sqs) (7) the mixing parameter (etas) and (8) the variance of the prior on alpha (psi_sqs). The list also contains the specified hyperparameters (hyperparam_a1, hyperparam_b1, hyperparam_a2, and hyperparam_b2), details regarding the dataset (X, C, Y, grp_idx), and Gibbs sampler details (n_burn_in, n_samples, and n_thin).

`rgig_use` *Function that calls the rgig function from the GeneralizedHyperbolic package in R.*

Description

Randomly generates one draw from a generalized inverse gaussian distribution.

Usage

```
rgig_use(chi, psi, lambda)
```

Arguments

`chi`	A positive double.
`psi`	A positive double.
`lambda`	A non-negative double.

Value

A random draw from the generalized inverse gaussian distribution with parameters chi, psi, and lambda (double).

Index