Package ‘BSGS’

February 19, 2015

Type Package
Title Bayesian Sparse Group Selection
Version 1.0
Date 2015-02-10
Author Kuo-Jung Lee and Ray-Bin Chen
Maintainer Kuo-Jung Lee <kuojunglee@mail.ncku.edu.tw>
Description The integration of Bayesian variable and sparse group variable selection approaches for regression models.
Depends R (>= 3.1.1), MASS, compiler, pscl, plyr
License GPL-2
LazyLoad yes
Repository CRAN
NeedsCompilation no
Date/Publication 2015-02-13 07:39:32

R topics documented:

BSGS ................................................................. 2
BSGS.Simple ..................................................... 3
CompWiseGibbsSimple ......................................... 4
CompWiseGibbsSMP .............................................. 5
Crisis2008 .......................................................... 6
Crisis2008BalancedData ....................................... 7

Index 8
BSGS

Sample version of group-wise Gibbs sampler for sparse group selection.

Description
Perform Bayesian sparse group selection to identify the important groups of variables and variables within those.

Usage
BSGS(y, X, Group.Index, r.value, eta.value, beta.value, tau2.value, rho.value, theta.value, sigma2.value, nu, lambda, Num.of.Iter.Inside.CompWise, Num.Of.Iteration, Burn.In)

Arguments
- **Y** vector of observations of length \( n \).
- **X** design matrix of dimension \( n \times p \).
- **Group.Index** Specify the group that each variable within design matrix belongs to.
- **r.value** Initial values of indicator variables for individual variables.
- **eta.value** Initial values of indicator variable for group variables.
- **beta.value** Initial values of regression coefficients, \( \beta \).
- **tau2.value** Variance in the prior distribution for regression coefficients.
- **rho.value** Prior probability including a variable.
- **theta.value** Prior probability including a group.
- **sigma2.value** Initial value of \( \sigma^2 \).
- **nu** Given value in the prior distribution of \( \sigma^2 \).
- **lambda** Given value in the prior distribution of \( \sigma^2 \).
- **Num.of.Iter.Inside.CompWise** Specify the number of iterations to be runned within component wise Gibbs sampler for variable selection within a group.
- **Num.Of.Iteration** Specify the number of iterations to be runned for sparse group variable selection.
- **Burn.In** Specify the number of iterations to be discarded.

Value
A list is returned with estimates of regression coefficients, \( \beta \), binary variables for group selection, \( \eta \), binary variables for variable selection, \( \gamma \), and variance, \( \sigma^2 \).
BSGS.Simple

Examples

```r
## Not run:
output = BSGS(Y, X, Group.Index, r.value, eta.value, beta.value, tau2.value, rho.value, theta.value, sigma2.value, nu, lambda, Num.of.Iter.Inside.CompWise, Num.Of.Iteration, Burn.In)
output$beta.est
output$eta.est
output$r.est
output$sigma2.est

## End(Not run)
```

BSGS.Simple  

*The group-wise Gibbs sampler for sparse group selection.*

Description

Perform Bayesian sparse group selection to identify the important groups of variables and variables within those.

Usage

```r
```

Arguments

<table>
<thead>
<tr>
<th>Argument</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Y</td>
<td>Vector of observations of length n.</td>
</tr>
<tr>
<td>X</td>
<td>Design matrix of dimension $n \times p$.</td>
</tr>
<tr>
<td>Group.Index</td>
<td>Specify the group that each variable within design matrix belongs to.</td>
</tr>
<tr>
<td>r.value</td>
<td>Initial values of indicator variables for individual variables.</td>
</tr>
<tr>
<td>eta.value</td>
<td>Initial values of indicator variable for group variables.</td>
</tr>
<tr>
<td>beta.value</td>
<td>Initial values of regression coefficients, $\beta$.</td>
</tr>
<tr>
<td>tau2.value</td>
<td>Variance in the prior distribution for regression coefficients.</td>
</tr>
<tr>
<td>rho.value</td>
<td>Prior probability including a variable.</td>
</tr>
<tr>
<td>theta.value</td>
<td>Prior probability including a group.</td>
</tr>
<tr>
<td>sigma2.value</td>
<td>Variance in the regression model.</td>
</tr>
<tr>
<td>nu</td>
<td>Given value in the prior distribution of $\sigma^2$.</td>
</tr>
<tr>
<td>lambda</td>
<td>Given value in the prior distribution of $\sigma^2$.</td>
</tr>
</tbody>
</table>
CompWiseGibbsSimple

Num.of.Iter.Inside.CompWise
   Specify the number of iterations to be runned within component wise Gibbs sampler for variable selection within a group.

Num.Of.Iteration
   Specify the number of iterations to be runned for sparse group variable selection.

Burn.In
   Specify the number of iterations to be discarded.

Value
   A list is returned with estimates of regression coefficients, $\beta$, binary variables for group selection, $\eta$, binary variables for variable selection, $r$, and variance, $\sigma^2$.

Examples

```r
## Not run:

output = bsgs.simple(y, x, group.index, r.value, eta.value, beta.value, tau2.value, rho.value, theta.value, sigma2.value, nu, lambda, Num.of.Iter.Inside.CompWise, Num.Of.Iteration, Burn.In)
output$beta.est
output$eta.est
output$r.est
output$sigma2.est

## End(Not run)
```

CompWiseGibbsSimple  Running MCMC procedure to generate the posterior samples from the posterior distribution.

Description

Perform Bayesian variable selection with component-wise prior to identify the important variables.

Usage

```
CompWiseGibbsSimple(Y, X, beta.value, r, tau2, rho, sigma2, nu, lambda, 
num.of.inner.iter, num.of.iteration)
```

Arguments

- **Y**: vector of observations of length $n$.
- **X**: design matrix of dimension $n \times p$.
- **beta.value**: Initial values of regression coefficients, $\beta$. 

CompWiseGibbsSimple  Running MCMC procedure to generate the posterior samples from the posterior distribution.
CompWiseGibbsSMP

\( r \)  
Initial values of indicator variables for individual regressors.

\( \tau^2 \)  
Variance in the prior distribution for regression coefficients.

\( \rho \)  
Prior probability including a variable.

\( \sigma^2 \)  
Initial value of \( \sigma^2 \).

\( \nu \)  
Given value in the prior distribution of \( \sigma^2 \).

\( \lambda \)  
Given value in the prior distribution of \( \sigma^2 \).

num.of.inner.iter
The number of iterations before sampling \( \sigma^2 \).

num.of.iteration
The number of iterations to be runned for sparse group variable selection.

Value
A list is returned with posterior samples of regression coefficients, \( \beta \), variance \( \sigma^2 \), and binary variables, \( \gamma \).

Examples

```r
## Not run: CompWiseGibbsSimple(Y, X, beta.value, r, tau2, rho, sigma2, nu, lambda,
num.of.inner.iter.default, num.of.iteration)
## End(Not run)
```

CompWiseGibbsSMP  
Stochastic matching pursuit for variable selection.

Description
Perform MCMC procedure to generate the posterior samples to estimate posterior quantities of interest in Bayesian variable selection using stochastic matching pursuit approach.

Usage

```
CompWiseGibbsSMP(Y, X, beta.value, r, tau2, rho, sigma2, nu, lambda,
num.of.inner.iter, num.of.iteration)
```

Arguments

\( Y \)  
vector of observations of length \( n \).

\( X \)  
design matrix of dimension \( n \times p \).

\( \text{beta.value} \)  
Initial values of regression coefficients, \( \beta \).

\( r \)  
Initial values of indicator variables for individual regressors.

\( \tau^2 \)  
Variance in the prior distribution for regression coefficients.
rho Prior probability including a variable.
sigma2 Initial value of $\sigma^2$.
nu Given value in the prior distribution of $\sigma^2$.
lambda Given value in the prior distribution of $\sigma^2$.
num.of.inner.iter The number of iterations before sampling $\sigma^2$.
num.of.iteration The number of iterations to be runned for sparse group variable selection.

Value
A list is returned with estimates of regression coefficients, $\beta$, binary variables for variable selection, $\gamma$, and variance, $\sigma^2$.

Examples

```r
## Not run: CompWiseGibbsSMP(Y, X, beta.value, r, tau2, rho, sigma2, nu0, lambda0, 
## num.of.inner.iter, num.of.iteration)
## End(Not run)
```

Crisis2008 A cross-sectional data set from Rose and Spiegel.

Description
A cross-sectional data set from Rose and Spiegel (2011), which is available at http://faculty.haas.berkeley.edu/arose. The response variable is 2008-2009 growth rate for the crisis measure. Rose and Spiegel originally consider 119 explanatory factors for the crisis for as many as 107 countries, but there are data missing for a number of countries.

Usage
data(Crisis2008)
Crisis2008BalancedData

A cross-sectional data set from Rose and Spiegel with the removal of missing values.

Description

A cross-sectional data set from Rose and Spiegel (2011), which is available at http://faculty.haas.berkeley.edu/arose. The response variable is 2008-2009 growth rate for the crisis measure. Rose and Spiegel originally consider 119 explanatory factors for the crisis for as many as 107 countries, but there are data missing for a number of countries. To maintain a balanced data set, we use 51 regressors for a sample of 72 countries. These regressors can be classified into the nine theoretical groups of the crisis’ origin (the number in parentheses indicates the number of variables considered in the group): principal factors (10), financial policies (three), financial conditions (four), asset price appreciation (two), macroeconomic policies (four), institutions (11), geography (four), financial linkages (one), and trade linkages (12).

Usage

data(Crisis2008BalancedData)
Index

BSGS, 2
BSGS.Simple, 3

CompWiseGibbsSimple, 4
CompWiseGibbsSMP, 5
Crisis2008, 6
Crisis2008BalancedData, 7