Package ‘mlsbm’

February 7, 2021

Title Efficient Estimation of Bayesian SBMs & MLSBMs

Version 0.99.2

Description Fit Bayesian stochastic block models (SBMs) and multi-level stochastic block models (MLSBMs) using efficient Gibbs sampling implemented in 'Rcpp'. The models assume symmetric, non-reflexive graphs (no self-loops) with unweighted, binary edges. Data are input as a symmetric binary adjacency matrix (SBMs), or list of such matrices (MLSBMs).

License GPL (>= 2)

Encoding UTF-8

LazyData true

RoxygenNote 7.1.1

LinkingTo Rcpp, RcppArmadillo

Imports Rcpp

Depends R (>= 2.10)

NeedsCompilation yes

Author Carter Allen [aut, cre] (<https://orcid.org/0000-0001-6937-7234>), Dongjun Chung [aut]

Maintainer Carter Allen <carter.allen12@gmail.com>

Repository CRAN

Date/Publication 2021-02-07 10:50:02 UTC

R topics documented:

<table>
<thead>
<tr>
<th>Topic</th>
<th>Pages</th>
</tr>
</thead>
<tbody>
<tr>
<td>AL</td>
<td>2</td>
</tr>
<tr>
<td>col_summarize</td>
<td>2</td>
</tr>
<tr>
<td>fit_mlsbm</td>
<td>3</td>
</tr>
<tr>
<td>fit_sbm</td>
<td>4</td>
</tr>
<tr>
<td>mean_CRI</td>
<td>5</td>
</tr>
<tr>
<td>mlsbm</td>
<td>5</td>
</tr>
<tr>
<td>sample_mlsbm</td>
<td>6</td>
</tr>
<tr>
<td>sample_sbm</td>
<td>6</td>
</tr>
</tbody>
</table>

Index 8
### AL

**Simulated 3-layer network data**

**Description**

A data set containing 3 layers of undirected, symmetric adjacency matrices simulated from an SBM with 3 true clusters

**Usage**

```
AL
```

**Format**

A list of length 3

---

### col_summarize

**The col_summarize function**

**Description**

Function to quickly return credible intervals

**Usage**

```
col_summarize(MAT, dig = 2, level = 0.95)
```

**Arguments**

- **MAT**: A matrix
- **dig**: Number of digits to round estimates and CrIs to
- **level**: Confidence level

**Value**

A character vector of posterior estimates and intervals

**Examples**

```
M <- matrix(rnorm(1000), ncol = 4)
col_summarize(M)
```
Description

This function allows you to fit multilevel stochastic block models.

Usage

```r
fit_mlsbm(
  A,
  K,
  a0 = 0.5,
  b10 = 0.5,
  b20 = 0.5,
  n_iter = 1000,
  burn = 100,
  verbose = TRUE
)
```

Arguments

- **A**: An adjacency list of length L, the number of levels. Each level contains an n x n symmetric adjacency matrix.
- **K**: The number of clusters specified a priori.
- **a0**: Dirichlet prior parameter for cluster sizes for clusters 1,...,K.
- **b10**: Beta distribution prior parameter for community connectivity.
- **b20**: Beta distribution prior parameter for community connectivity.
- **n_iter**: The number of total MCMC iterations to run.
- **burn**: The number of burn-in MCMC iterations to discard. The number of saved iterations will be n_iter - burn.
- **verbose**: Whether to print a progress bar to track MCMC progress. Defaults to true.

Value

A list of MCMC samples, including the MAP estimate of cluster indicators (z)

Examples

```r
data(AL)
# increase n_iter in practice
fit <- fit_mlsbm(AL,3,n_iter = 100)
```
fit_sbm  

R/Rcpp function for fitting single level stochastic block model

Description

This function allows you to fit single level stochastic block models.

Usage

fit_sbm(
  A,
  K,
  a0 = 0.5,
  b10 = 0.5,
  b20 = 0.5,
  n_iter = 1000,
  burn = 100,
  verbose = TRUE
)

Arguments

A  An n x n symmetric adjacency matrix.
K  The number of clusters specified a priori.
a0  Dirichlet prior parameter for cluster sizes for clusters 1,...,K.
b10  Beta distribution prior parameter for community connectivity.
b20  Beta distribution prior parameter for community connectivity.
n_iter  The number of total MCMC iterations to run.
burn  The number of burn-in MCMC iterations to discard. The number of saved iterations will be n_iter - burn.
verbose  Whether to print a progress bar to track MCMC progress. Defaults to true.

Value

A list of MCMC samples, including the MAP estimate of cluster indicators (z)

Examples

data(AL)
fit <- fit_sbm(AL[[1]],3)
The `mean_CRI` function

Description

Simple function to return the mean (95% CrI) for a vector

Usage

```r
mean_CRI(y, dig = 2)
```

Arguments

- `y`: A numeric vector
- `dig`: The number of digits to round to

Value

A string of mean and 95% quantile interval rounded to `dig`

Examples

```r
mean_CRI(rnorm(1000))
```

**mlsbm**

`mlsbm` functions

This package fits Bayesian stochastic block models (SBMs)

The `mlsbm` functions ...
**sample_mlsbm**

*R/Rcpp function for sampling from a multilevel stochastic block model*

**Description**

This function allows you to sample a multilevel stochastic block model.

**Usage**

```r
sample_mlsbm(z, P, L)
```

**Arguments**

- `z`: An n x 1 vector of community labels for each node
- `P`: A K x K symmetric matrix of community connectivity probabilities
- `L`: The number of levels to sample

**Value**

A list of adjacency matrices – one for each level of the MLSBM

**Examples**

```r
n = 100
K = 3
L = 2
pi = rep(1/K,K)
z = sample(1:K, size = n, replace = TRUE, prob = pi)
p_in = 0.50
p_out = 0.05
P = matrix(p_out, nrow = K, ncol = K)
diag(P) = p_in
AL = sample_mlsbm(z, P, L)
```

---

**sample_sbm**

*R/Rcpp function for sampling from a single level stochastic block model*

**Description**

This function allows you to sample a single level stochastic block model.

**Usage**

```r
sample_sbm(z, P)
```

**Examples**

```r
n = 100
K = 3
z = sample(1:K, size = n, replace = TRUE, prob = pi)
p_in = 0.50
p_out = 0.05
P = matrix(p_out, nrow = K, ncol = K)
diag(P) = p_in
AL = sample_mlsbm(z, P)
```
sample_sbm

Arguments

* z 
  An n x 1 vector of community labels for each node
* P 
  A K x K symmetric matrix of community connectivity probabilities

Value

An adjacency matrix

Examples

```r
n = 100
K = 3
pi = rep(1/K, K)
z = sample(1:K, size = n, replace = TRUE, prob = pi)
p_in = 0.50
p_out = 0.05
P = matrix(p_out, nrow = K, ncol = K)
diag(P) = p_in
A = sample_sbm(z, P)
```
Index

* Bayesian
  fit_mlsbm, 3
  fit_sbm, 4
* Credible
  col_summarize, 2
* Gibbs
  fit_mlsbm, 3
  fit_sbm, 4
* MCMC
  mean_CRI, 5
* MLSBM
  fit_mlsbm, 3
  fit_sbm, 4
* SBM
  fit_mlsbm, 3
  fit_sbm, 4
  sample_mlsbm, 6
  sample_sbm, 6
* datasets
  AL, 2
* intervals
  col_summarize, 2
* networks
  fit_mlsbm, 3
  fit_sbm, 4

AL, 2

col_summarize, 2

fit_mlsbm, 3
fit_sbm, 4

mean_CRI, 5
mlsbm, 5

sample_mlsbm, 6
sample_sbm, 6