Package ‘lglasso’

October 13, 2022

Type Package

Title Longitudinal Graphical Lasso

Version 0.1.0

Description For high-dimensional correlated observations, this package carries out the L_1 penalized maximum likelihood estimation of the precision matrix (network) and the correlation parameters. The correlated data can be longitudinal data (may be irregularly spaced) with dampening correlation or clustered data with uniform correlation. For the details of the algorithms, please see the paper Jie Zhou et al. Identifying Microbial Interaction Networks Based on Irregularly Spaced Longitudinal 16S rRNA sequence data <doi:10.1101/2021.11.26.470159>.

License GPL-3

Encoding UTF-8

LazyData true

RoxygenNote 7.1.2

URL https://github.com/jiezhou-2/lglasso

Suggests knitr, rmarkdown

Imports stats, glasso

Depends R (>= 2.10)

NeedsCompilation no

Author Jie Zhou [aut, cre, cph], Jiang Gui [aut], Weston Viles [aut], Anne Hoen [aut]

Maintainer Jie Zhou <chowstat@gmail.com>

Repository CRAN

Date/Publication 2022-01-15 08:52:46 UTC
heterlongraph

**R topics documented:**

- heterlongraph
- homolongraph
- iss
- lglasso
- lii_homo
- ll_homo
- logdensity
- mle
- mle_alpha
- mle_net
- mle_tau
- phifunction
- sample_data

**Index**

heterlongraph 13

---

**Description**

Estimates of correlation parameters and precision matrix

**Usage**

```r
heterlongraph(data, rho, type, tole, lower, upper)
```

**Arguments**

- **data** Data matrix in which the first column is subject id, the second column is the time points of observation. Columns 2 to (p+2) is the observations for p variables.
- **rho** Tuning parameter used in graphical lasso
- **type** Type of correlation function, which can take either "abs" or "sqr".
- **tole** Error tolerance for determination of convergence of EM algorithm
- **lower** Lower bound for prediction of correlation parameter tau
- **upper** Upper bound for prediction of correlation parameter tau

**Value**

S list with three components which are the final estimate of alpha, tau and precision matrix omega

**Author(s)**

Jie Zhou
homolongraph

Estimation of precision matrix and autocorrelation parameter for homogeneous model

Description

Estimation of precision matrix and autocorrelation parameter for homogeneous model

Usage

homolongraph(data, rho, type, tole, lower, upper)

Arguments

data Data matrix in which the first column is subject ID, the second column is the time points of observation. Columns 2 to (p+2) is the observations for p variables.
rho Tuning parameter for graphical lasso
type Type of correlation function, which can take either "abs" or "qua".
tole Error tolerance for determination of convergence of EM algorithm
lower Lower bound for prediction of correlation parameter tau
upper Upper bound for prediction of correlation parameter tau

Value

A list for estimates of precision matrix and correlation parameter for given tuning parameter

Author(s)

Jie Zhou

iss

Quasi covariance matrix for subject i

Description

Quasi covariance matrix for subject i

Usage

iss(idata, itau, type)
Arguments

idata  Data matrix for the subject i in which the first column is subject (cluster) id, the second column stands for the time points () of observation. Columns 2 to (p+2) is the observations for p variables respectively.

itau  Correlation parameter

type  Type of correlation function, which typically take either 0, 1 or 2.

Value

Empirical quasi covariance matrix

Author(s)

Jie Zhou

Description

This function implements the $L_1$ penalized maximum likelihood estimation for precision matrix (network) based on correlated data, e.g., irregularly spaced longitudinal data. It can be regarded as an extension of the package glasso (Friedman, Hastie and Tibshirani, 2008) which aims to find the sparse estimate of the network from independent continuous data.

Usage

```r
lglasso(
  data,
  rho,
  heter = TRUE,
  type = 1,
  tole = 0.01,
  lower = 0.01,
  upper = 10
)
```

Arguments

- **data**: Data matrix in which the first column is subject id, the second column is time points of observations for temporal data or site id for spatial data. Columns 3 to (p+2) is the observations for p variables.
- **rho**: Tuning parameter used in $L_1$ penalty
**heter**  Binary variable TRUE or FALSE, indicating heterogeneous model or homogeneous model is fitted. In heterogeneous model, subjects are allowed to have his/her own temporal correlation parameter \( \tau_i \); while in homogeneous model, all the subjects are assumed to share the same temporal correlation parameter, i.e., \( \tau_1 = \tau_2 = \ldots = \tau_m \).

**type**  A positive number which specify the correlation function. The general form of correlation function is given by \( \exp(\tau|t_i-t_j|^\text{type}) \). in which \( \text{type}=0 \) can be used for spatial correlation while \( \text{type}>0 \) are used for temporal correlation. For latter, the default value is set to be \( \text{type}=1 \).

**tole**  Threshold for convergence. Default value is \( 1e^{-2} \). Iterations stop when maximum absolute difference between consecutive estimates of parameter change is less than tole.

**lower**  Lower bound for predicts of correlation parameter \( \tau \). Default value is \( 1e^{-2} \). The estimate of \( \tau(\alpha) \) will be searched in the interval \([\text{lower}, \text{upper}]\), where parameter upper is explained in the following.

**upper**  Upper bound for predicts of correlation parameter \( \tau \).

**Value**

If heter=TRUE, then a list with three components is returned which are respectively the estimate of parameter alpha in exponent distribution, correlation parameter tau and precision matrix omega. If heter=FALSE, then a list with two components is returned which are respectively the estimate of correlation parameter tau and precision matrix omega.

**Author(s)**

Jie Zhou

**References**


**Examples**

```r
sample_data[1:5,1:5]
dim(sample_data)
## Heterogeneous model with dampening correlation rate using the first three clusters
a=lglasso(data = sample_data[1:11,], rho = 0.7,heter=TRUE, type=1)
### Estimates of correlation parameters
a$tau
### Sub-network for the first five variables
a$omega[1:5,1:5]
```
### Total number of the edges in the estimated network
(length(which(a$omega!=0))-ncol(a$omega))/2

## Homogeneous model with dampening correlation rate using the first three clusters
b=lglasso(data = sample_data[1:11,], rho = 0.7,heter=FALSE,type=1)
### Estimates of correlation parameters
b$tau
### Sub-network for the first five variables
b$omega[1:5,1:5]
### Total number of the edges in the estimated network
(length(which(b$omega!=0))-ncol(b$omega))/2

## Heterogeneous model with uniform correlation rate using the first three clusters
c=lglasso(data = sample_data[1:11,], rho = 0.7,heter=TRUE,type=0)
### Estimates of correlation parameters
c$tau
### Sub-network for the first five variables
c$omega[1:5,1:5]
### Total number of the edges in the estimated network
(length(which(c$omega!=0))-ncol(c$omega))/2

## Homogeneous model with uniform correlation rate using the first three clusters
d=lglasso(data = sample_data[1:11,], rho = 0.7,heter=FALSE,type=0)
### Estimates of correlation parameters
d$tau
### Sub-network for the first five variables
d$omega[1:5,1:5]
### Total number of the edges in the estimated network
(length(which(d$omega!=0))-ncol(d$omega))/2

---

**lli_homo**

*full log likelihood used in EBIC computation*

**Description**

full log likelihood used in EBIC computation

**Usage**

lli_homo(idata, omega, tau, type)

**Arguments**

<table>
<thead>
<tr>
<th>Argument</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>idata</td>
<td>Data matrix for the subject i in which the first column is id for subject, the second column is the time points of observation. Columns 2 to (p+2) is the observations for p variables.</td>
</tr>
<tr>
<td>omega</td>
<td>Precision matrix</td>
</tr>
<tr>
<td>tau</td>
<td>Correlation parameter</td>
</tr>
<tr>
<td>type</td>
<td>Type of correlation function, which can take either &quot;abs&quot; or &quot;qua&quot;.</td>
</tr>
</tbody>
</table>
### Description

Value of likelihood function at given parameter

### Usage

```r
ll_homo(data, omega, tau, type)
```

### Arguments

- **data**: Data matrix in which the first column is subject id, the second column is the time points of observation. Columns 2 to (p+2) is the observations for p variables.
- **omega**: Precision matrix
- **tau**: Correlation parameter
- **type**: Type of correlation function, which can take either "abs" or "qua".

### Value

Value of likelihood function at given omega and tau

### Author(s)

Jie Zhou
logdensity

*Complete likelihood function used in EM algorithm of heterogeneous marginal graphical lasso model*

**Description**

Complete likelihood function used in EM algorithm of heterogeneous marginal graphical lasso model

**Usage**

```
logdensity(idata, omega, tau, alpha, type)
```

**Arguments**

- **idata**: Data matrix for the subject i in which the first column is id for subject, the second column is the time points of observation. Columns 2 to (p+2) is the observations for p variables.
- **omega**: Precision matrix
- **tau**: Correlation parameter
- **alpha**: Parameter in exponential distribution
- **type**: Type of correlation function, which can take either "abs" or "qua".

**Value**

Value of complete likelihood function at given value of omega, tau and alpha

**Author(s)**

Jie Zhou

---

mle

*Maximum Likelihood Estimate of Precision Matrix and Correlation Parameters for Given Network*

**Description**

Maximum Likelihood Estimate of Precision Matrix and Correlation Parameters for Given Network
Usage

mle(
  data,
  network,
  heter = TRUE,
  type = 1,
  tole = 0.01,
  lower = 0.01,
  upper = 10
)

Arguments

data Data matrix in which the first column is subject id, the second column is time points of observations for temporal data or site id for spatial data. Columns 3 to (p+2) is the observations for p variables.

network The network selected by function lglasso

heter Binary variable TRUE or FALSE, indicating heterogeneous model or homogeneous model is fitted. In heterogeneous model, subjects are allowed to have his/her own temporal correlation parameter $\tau_i$; while in homogeneous model, all the subjects are assumed to share the same temporal correlation parameter, i.e., $\tau_1=\tau_2=\ldots=\tau_m$.

type A positive number which specify the correlation function. The general form of correlation function is given by $\exp(|t_i-t_j|^\text{type})$, in which $\text{type}=0$ can be used for spatial correlation while $\text{type}>0$ are used for temporal correlation. For latter, the default value is set to be $\text{type}=1$.

tole Threshold for convergence. Default value is $1e^{-2}$. Iterations stop when maximum absolute difference between consecutive estimates of parameter change is less than tole.

lower Lower bound for predicts of correlation parameter tau. Default value is $1e^{-2}$. The estimate of $\tau(\alpha)$ will be searched in the interval $[\text{lower}, \text{upper}]$, where parameter upper is explained in the following.

upper Upper bound for predicts of correlation parameter tau.

Value

A list which include the maximum likelihood estimate of precision matrix, correlation parameter $\tau$. If heter=TRUE, the output also include the estimate of alpha where $\tau=\exp(\alpha)$

Author(s)

Jie Zhou
mle_alpha  

**Maximum likelihood estimate of correlation parameter for given structure of precision matrix**

**Description**

Maximum likelihood estimate of correlation parameter for given structure of precision matrix

**Usage**

```r
mle_alpha(data, alpha0, omega, type, tole, lower, upper)
```

**Arguments**

- **data**: Data matrix in which the first column is subject id, the second column is the time points of observation. Columns 2 to (p+2) is the observations for p variables.
- **alpha0**: Initial value for the parameter in exponential distribution
- **omega**: Fixed value for precision matrix
- **type**: Type of correlation function, which can take either "abs" or "qua".
- **tole**: Error tolerance for determination of convergence of EM algorithm
- **lower**: Lower bound for prediction of correlation parameter tau
- **upper**: Upper bound for prediction of correlation parameter tau

**Author(s)**

Jie Zhou

---

mle_net  

**Title**

**Description**

Title

**Usage**

```r
mle_net(data, priori)
```

**Arguments**

- **data**: A Longitudinal data set
- **priori**: Given structure of precision matrix
Value

The maximum likelihood estimation

Author(s)

Jie Zhou

---

### mle_tau

*Estimate of precision matrix and autocorrelation parameter for homogeneous model*

---

**Description**

Estimate of precision matrix and autocorrelation parameter for homogeneous model

**Usage**

```r
mle_tau(data, omega, type, lower, upper)
```

**Arguments**

- `data`: Data matrix in which the first column is subject id, the second column is the time points of observation. Columns 2 to (p+2) is the observations for p variables.
- `omega`: The maximum likelihood estimate of precision matrix
- `type`: Type of correlation function, which can take either "abs" or "qua".
- `lower`: Lower bound for prediction of correlation parameter tau
- `upper`: Upper bound for prediction of correlation parameter tau

**Value**

A list for estimates of precision matrix and correlation parameter for given tuning parameter

**Author(s)**

Jie Zhou
phifunction  

Construct the temporal component for correlation function

Description

Construct the temporal component for correlation function

Usage

phifunction(t, tau, type = 1)

Arguments

t  Time points of observations
tau  correlation parameter
type  The type of correlation function, which typically take either 0, 1 or 2.

Value

A square matrix with dimension equal to the length of vector t

Author(s)

Jie Zhou

sample_data  

Sample Data

Description

The sample data are subset of a larger longitudinal data set from an ongoing large-scale prospective project. There are 13 cluster are involved in the sample data.

Usage

sample_data

Format

A 100-by-22 matrix

Column 1  Cluster id;
Column 2  Time points of observations;
Columns 3-22  Observations for 20 microbes.
**Index**

* datasets
  - sample_data, 12

heterlongraph, 2
homolongraph, 3

iss, 3

lglasso, 4
ll_homo, 7
lli_homo, 6
logdensity, 8

mle, 8
mle_alpha, 10
mle_net, 10
mle_tau, 11

phifunction, 12

sample_data, 12