

Package ‘SparseTSCGM’

August 21, 2014

Type Package

Title Sparse time series chain graphical models

Version 2.1.1

Date 2014-04-15

Depends R(>= 3.0.2)

Imports glasso, longitudinal, flare, MASS, mvtnorm, network, abind

Author Fentaw Abegaz and Ernst Wit

Maintainer Fentaw Abegaz <f.abegaz.yazew@rug.nl>

Description Computes sparse vector autoregressive coefficients and precision matrices for time series chain graphical models.

License GPL (>= 3)

LazyLoad yes

NeedsCompilation yes

Repository CRAN

Date/Publication 2014-08-21 06:43:12

R topics documented:

SparseTSCGM-package	2
mammary	3
plot.tscgm	4
plot.tscgm.ar2	5
print.tscgm	6
sim.data	7
sparse.tscgm	8
summary.tscgm	11

Index	13
--------------	-----------

SparseTSCGM-package *Sparse time series chain graphical models.*

Description

Computes sparse autoregressive coefficient and precision matrices for time series chain graphical models (TSCGM). These models provide an efficient way of simultaneously dealing with Gaussian graphical models (undirected graphs for instantaneous interactions) and Bayesian networks (directed graphs for dynamic interactions) for reconstructing instantaneous and dynamic networks from repeated multivariate time series data.

Details

Package:	SparseTSCGM
Type:	Package
Version:	2.0
Date:	2014-01-31
License:	GPL (>=3)
LazyLoad:	yes

Author(s)

Fentaw Abegaz and Ernst Wit

Maintainer: Fentaw Abegaz <f.abegaz.yazew@rug.nl>

References

Fentaw Abegaz and Ernst Wit (2013). *Sparse time series chain graphical models for reconstructing genetic networks*. *Biostatistics*. **14**, 3: 586-599.

Rothman, A.J., Levina, E., and Zhu, J. (2010). *Sparse multivariate regression with covariance estimation*. *Journal of Computational and Graphical Statistics*. **19**: 947-962.

Examples

```
set.seed(321)
datas <- sim.data(model="ar1", time=20, n.obs=25, n.var=10, prob0=0.35,
  network="random")
data.fit <- datas$data1
prec_true <- datas$theta
autoR_true <- datas$gamma
```

```
res.tscgm <- sparse.tscgm(data=data.fit, lam1=NULL, lam2=NULL, nlambda=NULL,
  model="ar1", optimality="bic_ext",control=list(maxit.out = 10, maxit.in = 100))

#Estimated sparse precision and autoregression matrices
prec <- res.tscgm$theta
autoR <- res.tscgm$gamma

#Graphical visualization
par(mfrow=c(2,2))
plot.tscgm(datas, mat="precision",main="True precision matrix")
plot.tscgm(res.tscgm, mat="precision",main="Estimated precision matrix")
plot.tscgm(datas, mat="autoregression",main="True autoregression coef. matrix")
plot.tscgm(res.tscgm, mat="autoregression",
  main="Estimated autoregression coef. matrix")
```

mammary

Microarray gene expression time course data for mammary gland development in mice

Description

The data contains 30 genes identified using cluster analysis from 12488 probe sets representing approximately 8600 genes. The data are measured over 54 arrays of 3 replicates each on 18 time points of the developmental stages of mammary gland in mice. See Stein et al. (2004) for more details.

Usage

```
data(mammary)
```

Format

Data is in longitudinal format with 30 columns, 54 rows and a number of extra attributes (see R package longitudinal).

Source

This data is described in Stein et al. (2004) and can be freely obtained from the R package smida.

References

Stein T, Morris J, Davies C, Weber Hall S, Duffy M, Heath V, Bell A, Ferrier R, Sandilands G, Gusterson B, et al. (2004). Involution of the mouse mammary gland is associated with an immune cascade and an acute phase response, involving LBP, CD14 and STAT3. *Breast Cancer Res*, 6(2), R 75 - 91.

Wit E. and McClure J. (2004). *Statistics for Microarrays: Design, Analysis and Inference*. Wiley.

Examples

```
# load "longitudinal" library
library(longitudinal)

# load data sets
data(mammary)
```

plot.tscgm	<i>Plot sparse.tscgm objects from fitting chain graphical models with vector autoregressive process of order 2.</i>
------------	---

Description

plot.tscgm is a generic plot function that is adapted for objects of class sparse.tscgm.

Usage

```
## S3 method for class 'tscgm'
plot(x, mat=c("precision", "autoregression"), ...)
```

Arguments

x	an object of class sparse.tscgm.
mat	Name of matrix to be plotted, i.e., either the precision matrix or vector autoregression matrix.
...	Arguments to be passed to graphical parameters (see par).

Value

Undirected or directed networks.

Author(s)

Fentaw Abegaz and Ernst Wit

See Also

[network](#)

Examples

```
set.seed(321)
datas <- sim.data(model="ar1", time=20, n.obs=25, n.var=10, prob0=0.35,
                 network="random")
data.fit <- datas$data1

res.tscgm <- sparse.tscgm(data=data.fit, lam1=NULL, lam2=NULL, nlambdas=NULL,
```

```

model="ar1", optimality="bic_ext",control=list(maxit.out = 10, maxit.in = 100))

#Network visualization
par(mfrow=c(2,1))
plot.tscgm(res.tscgm, mat="precision", main="Undirected network", pad = 0.01,
  label.pad = 0.3, label.col = 6, vertex.col = 5,vertex.cex = 1.5,
  edge.col = 4, mode = "fruchtermanreingold", interactive=FALSE)

plot.tscgm(res.tscgm, mat="autoregression", main="Directed network", pad = 0.01,
  label.pad = 0.3, label.col = 6, vertex.col = 5,vertex.cex = 1.5,
  edge.col = 4, mode = "fruchtermanreingold", interactive=FALSE)

```

plot.tscgm.ar2	<i>Plot sparse.tscgm objects from fitting chain graphical models with vector autoregressive process of order 2.</i>
----------------	---

Description

plot.tscgm.ar2 is a generic plot function that is adapted for objects of class sparse.tscgm.

Usage

```

## S3 method for class 'tscgm.ar2'
plot(x, mat=c("precision", "autoregression1", "autoregression2"),...)

```

Arguments

x	an object of class sparse.tscgm.
mat	Name of matrix to be plotted,i.e., either the precision matrix or vector autoregression matrices of lag 1 or 2.
...	Arguments to be passed to graphical parameters (see par).

Value

Undirected or directed networks.

Author(s)

Fentaw Abegaz and Ernst Wit

See Also

[network](#)

Examples

```

## Data generation from time series chain graphical model with vector
## autoregressive model of order 2
set.seed(321)
datas <- sim.data(model="ar2", time=20,n.obs=50, n.var=10,prob0=0.25,
                 network="random")
data.fit <- datas$data1

## Model fitting with vector autoregressive order 2
res.tscgm <- sparse.tscgm(data=data.fit, lam1=NULL, lam2=NULL, nlambda=NULL,
                        model="ar2", optimality="bic_mod",control=list(maxit.out = 10, maxit.in = 100))

#Network visualization
par(mfrow=c(3,2))

#Graphical visualization
par(mfrow=c(3,2))
plot.tscgm.ar2(datas, mat="precision",main="True precision matrix")
plot.tscgm.ar2(res.tscgm, mat="precision",main="Estimated precision matrix")

plot.tscgm.ar2(datas, mat="autoregression1",
               main="True autoregression coef. matrix of lag 1" )
plot.tscgm.ar2(res.tscgm, mat="autoregression1",
               main="Estimated autoregression coef. matrix of lag 1")

plot.tscgm.ar2(datas, mat="autoregression2",
               main="True autoregression coef. matrix of lag 2")
plot.tscgm.ar2(res.tscgm, mat="autoregression2",
               main="Estimated autoregression coef. matrix of lag 2")

```

```
print.tscgm
```

```
Print summary - S3 Method for Class 'sparse.tscgm'.
```

Description

print is a generic function that prints output summaries of fitted models in the SparseTSCGM package.

Usage

```
## S3 method for class 'tscgm'
print(x, ...)
```

Arguments

x an object of class sparse.tscgm
... other arguments passed to print.

Details

The `print.tscgm` function summarizes and prints results of the fitted model.

Value

Prints summary of results:

Author(s)

Fentaw Abegaz and Ernst Wit

See Also

[sparse.tscgm](#)

Examples

```
# print.tscgm(res.tscgm)
```

sim.data

Multivariate time series simulation with chain graphical models

Description

Generates sparse vector autoregressive coefficients matrices and precision matrix from various network structures and using these matrices generates repeated multivariate time series dataset.

Usage

```
sim.data(model=c("ar1", "ar2"), time=time, n.obs=n.obs, n.var=n.var, prob0=NULL,
         network=c("random", "scale-free", "hub", "user_defined"),
         prec=NULL, gamma1=NULL, gamma2=NULL)
```

Arguments

<code>model</code>	Specifies the order of vector autoregressive models. Vector autoregressive model of order 1 is applied if <code>model = "ar1"</code> and Vector autoregressive model of order 2 is applied if <code>method = "ar2"</code> .
<code>time</code>	Number of time points.
<code>n.obs</code>	Number of observations or replicates.
<code>n.var</code>	Number of variables.
<code>prob0</code>	Initial sparsity level.
<code>network</code>	Specifies the type of network structure. This could be random, scale-free, hub or user defined structures. Details on simulations from the various network structures can be found in the R package <code>flare</code> .

prec	Precision matrix.
gamma1	Autoregressive coefficients matrix at time lag 1.
gamma2	Autoregressive coefficients matrix at time lag 2.

Value

A list containing:

theta	Sparse precision matrix.
gamma	Sparse autoregressive coefficients matrix.
data1	Repeated multivariate time series data in longitudinal format.

Author(s)

Fentaw Abegaz and Ernst Wit

Examples

```
set.seed(321)
datas <- sim.data(model="ar1", time=20,n.obs=20, n.var=20,prob0=0.15,
                 network="hub")
data.ts <- datas$data1
prec_true <- datas$theta
autoR_true <- datas$gamma
```

sparse.tscgm

Sparse time series chain graphical models

Description

Computes sparse vector autoregressive coefficients matrices of order 1 and order 2 and precision matrix estimates for time series chain graphical models using SCAD penalty. In time series chain graphs directed edges are identified by nonzero entries of the autoregressive coefficients matrix and undirected edges are identified by nonzero entries of the precision matrix.

Usage

```
sparse.tscgm(data = data, lam1 = NULL, lam2 = NULL, nlambda = NULL,
             model = c("ar1", "ar2"),
             optimality = c("NULL", "bic", "bic_ext", "bic_mod", "aic", "gic"),
             control = list())
```


Arguments

data	Longitudinal data format.
lam1	Either a scalar or a vector of tuning parameter values for the SCAD penalty on the precision matrix. The default is lam1=NULL and the program generates a sequence of tuning parameter values based on nlambda. The user can also specify a sequence to change the default sequence. If lam1 is a vector with length at least 2, information criteria based model selection will be performed to select the optimal tuning parameter.
lam2	Either a scalar or a vector of tuning parameter values for the SCAD penalty on the precision matrix. The default is lam2=NULL and the program generates a sequence of tuning parameter values based on nlambda. The user can also specify a sequence to change the default sequence. If lam2 is a vector with length at least 2, information criteria based model selection will be performed to select the optimal tuning parameter.
nlambda	The number of values used in lam1 and lam2. Default value is 10.
model	This specifies the order of vector autoregressive models. Vector autoregressive model of order 1 is applied if model = "ar1" and Vector autoregressive model of order 2 is applied if method = "ar2".
optimality	This specifies the type of information based model selection criteria. When optimality is "NULL" it computes the time series chain graphical model solutions for specified scalar values of the tuning parameters lam1 and lam2. The rest optimality criteria first selects the optimal SCAD tuning parameter values using the minimum information criterion from vectors of tuning parameters for lam1 and lam2 then computes the time series chain graphical model solutions based on the optimal tuning parameter values. These are: "BIC" (Bayesian Information Criteria), "bic_ext" (Extended Bayesian Information Criteria), "bic_mod" (Modified Bayesian Information Criteria) "aic" (Akaike Information Criteria), and "gic" (Generalized Information Criteria).
control	The argument control includes a list of control parameters passed to the main function. The user can specify its own control parameter values to replace the default values. <pre>control = list(maxit.out = 5, maxit.in = 50, tol.out = 1e-04, silent = TRUE)</pre> <p>maxit.out : The maximum allowable outer iterations of the algorithm. The outer-loop iterates alternatively to estimate the precision matrix and the autoregression matrix until convergence. Default is 5. maxit.in: The maximum allowable iterations of the algorithm that minimizes with respect to the autoregression coefficient matrix. Default is 50. tol.out: Convergence tolerance for outer-loop of the algorithm that alternate between the estimation of precision matrix and autoregressive coefficients matrix. Default is 1e-04. silent: Either TRUE or FALSE. If silent=FALSE this function will display progress updates to the screen. Default is TRUE.</p>

Details

For description of the objective functions and computational details see Abegaz and Wit (2013).

Value

A list containing:

theta	Precision matrix estimate. The nonzero estimates of the precision matrix are used for constructing undirected graphs that represent conditional independences at time lag 0 or instantaneous interactions.
gamma	Autoregressive coefficients matrix estimate. The nonzero estimates of the autoregression matrix are used for constructing directed graphs that represent conditional independences at time lag 1, time lag 2 or dynamic interactions.
lam1.opt	The optimal tuning parameter for SCAD penalty on the precision matrix selected with minimum information criterion.
lam2.opt	The optimal tuning parameter for SCAD penalty on the autoregressive coefficients matrix selected with minimum BIC criterion.
min.ic	Minimum value of the optimality criterion.
tun.ic	A matrix containing the list of tuning parameter values and the corresponding model selection or optimality values.
lam1.seq	The sequence of tuning parameter values related to precision matrix.
lam2.seq	The sequence of tuning parameter values related to autoregression matrix.
s.theta	Sparsity level of the precision matrix.
s.gamma	Sparsity level of the autoregression coefficients matrix.

Author(s)

Fentaw Abegaz and Ernst Wit

References

Fentaw Abegaz and Ernst Wit (2013). Sparse time series chain graphical models for reconstructing genetic networks. Biostatistics. 14, 3: 586-599.

Rothman, A.J., Levina, E., and Zhu, J. (2010). Sparse multivariate regression with covariance estimation. Journal of Computational and Graphical Statistics. 19: 947-962.

Examples

```
set.seed(321)
datas <- sim.data(model="ar1", time=20,n.obs=25, n.var=10,prob0=0.35,
                 network="random")
data.fit <- datas$data1
prec_true <- datas$theta
autoR_true <- datas$gamma

res.tscgm <- sparse.tscgm(data=data.fit, lam1=NULL, lam2=NULL, nlambda=NULL,
                        model="ar1", optimality="bic_ext",control=list(maxit.out = 10, maxit.in = 100))

#Estimated sparse precision and autoregression matrices
```

```

prec <- res.tscgm$theta
autoR <- res.tscgm$gamma

#Optimal tuning parameter values
lambda1.opt <- res.tscgm$lam1.opt
lambda2.opt <- res.tscgm$lam2.opt

#Sparsity levels
sparsity_theta <- res.tscgm$s.theta
sparsity_gamma <- res.tscgm$s.gamma

#Graphical visualization
par(mfrow=c(2,2))
plot.tscgm(datas, mat="precision",main="True precision matrix")
plot.tscgm(res.tscgm, mat="precision",main="Estimated precision matrix")
plot.tscgm(datas, mat="autoregression",main="True autoregression coef. matrix")
plot.tscgm(res.tscgm, mat="autoregression",
           main="Estimated autoregression coef. matrix")

```

summary.tscgm

Summary - S3 Method for Class 'sparse.tscgm'

Description

summary is a generic function that produces output summaries of fitted models in the SparseTSCGM package. In particular, the function invokes methods for objects of class [sparse.tscgm](#).

Usage

```

## S3 method for class 'tscgm'
summary(object, ...)

```

Arguments

object	An object of class <code>sparse.tscgm</code>
...	Other arguments passed to <code>summary</code> .

Details

The `summary.stscgm` function provides summary results of the fitted model.

Value

A list containing:

theta	Precision matrix estimate
gamma	Autoregressive coefficients matrix estimate

lam1.opt	The optimal tuning parameter on the precision matrix with model selection.
lam2.opt	The optimal tuning parameter on the autoregressive coefficients matrix with model selection.
min.ic	Minimum value of the optimality criterion.
tun.ic	A matrix containing the list values from model selection.
lam1.seq	The sequence of tuning parameter values on precision matrix.
lam2.seq	The sequence of tuning parameter values on autoregression matrix.
s.theta	Sparsity level of the precision matrix.
s.gamma	Sparsity level of the autoregression coefficients matrix.

Author(s)

Fentaw Abegaz and Ernst Wit

See Also

[sparse.tscgm](#)

Examples

```
# summary.tscgm(res.tscgm)
```

Index

*Topic **package**

SparseTSCGM-package, [2](#)

mammary, [3](#)

network, [4](#), [5](#)

par, [4](#), [5](#)

plot.tscgm, [4](#)

plot.tscgm.ar2, [5](#)

print.tscgm, [6](#)

sim.data, [7](#)

sparse.tscgm, [7](#), [8](#), [11](#), [12](#)

SparseTSCGM (SparseTSCGM-package), [2](#)

SparseTSCGM-package, [2](#)

summary.tscgm, [11](#)