Package ‘IsingFit’

September 7, 2016

Type Package

Title Fitting Ising Models Using the ELasso Method

Version 0.3.1

Date 2016-9-6

Depends R (>= 3.0.0)

Imports qgraph, Matrix, glmnet

Suggests IsingSampler

Author Claudia van Borkulo, Sacha Epskamp, with contributions from Alexander Robitzsch

Maintainer Claudia van Borkulo <cvborkulo@gmail.com>

Description This network estimation procedure eLasso, which is based on the Ising model, combines l1-regularized logistic regression with model selection based on the Extended Bayesian Information Criterion (EBIC). EBIC is a fit measure that identifies relevant relationships between variables. The resulting network consists of variables as nodes and relevant relationships as edges. Can deal with binary data.

License GPL-2

NeedsCompilation no

Repository CRAN

Date/Publication 2016-09-07 13:01:58

R topics documented:

`IsingFit-package` *Network estimation using the eLasso method*

Description

This network estimation procedure eLasso, which is based on the Ising model, combines l1-regularized logistic regression with model selection based on the Extended Bayesian Information Criterion (EBIC). EBIC is a fit measure that identifies relevant relationships between variables. The resulting network consists of variables as nodes and relevant relationships as edges. Can deal with binary data.

Details

Package:	IsingFit
Type:	Package
Version:	0.3.1
Date:	2016-9-6
License:	What license is it under?

Author(s)

Claudia D. van Borkulo, Sacha Epskamp, with contributions from Alexander Robitzsch

Maintainer: Claudia D. van Borkulo <cvborkulo@gmail.com>

References

Chen, J., & Chen, Z. (2008). Extended bayesian information criteria for model selection with large model spaces. Biometrika, 95(3), 759-771.

Foygel, R., & Drton, M. (2011). Bayesian model choice and information criteria in sparse generalized linear models. arXiv preprint arXiv:1112.5635.

Ravikumar, P., Wainwright, M. J., & Lafferty, J. D. (2010). High-dimensional Ising model selection using l1-regularized logistic regression. The Annals of Statistics, 38, 1287 - 1319.

van Borkulo, C. D., Borsboom, D., Epskamp, S., Blanken, T. F., Boschloo, L., Schoevers, R. A., & Waldorp, L. J. (2014). A new method for constructing networks from binary data. Scientific Reports 4, 5918; DOI:10.1038/srep05918.

Ising-methods *Methods for IsingFit objects*

Description

Print method prints the IsingFit output , plot method plots the estimated network (with the qgraph package), and summary method returns density of the network, the value of gamma used, the rule used, and the time the analysis took.

Usage

```
## S3 method for class 'IsingFit'
print(x, ...)
## S3 method for class 'IsingFit'
summary(object, ...)
## S3 method for class 'IsingFit'
plot(x, ...)
```

Arguments

x	output of IsingFit
object	output of IsingFit
...	Arguments sent to qgraph. Only used in plot method.

Author(s)

Claudia van Borkulo

IsingFit *Network estimation using the eLasso method*

Description

This network estimation procedure eLasso, which is based on the Ising model, combines l1-regularized logistic regression with model selection based on the Extended Bayesian Information Criterion (EBIC). EBIC is a fit measure that identifies relevant relationships between variables. The resulting network consists of variables as nodes and relevant relationships as edges. Can deal with binary data.

Usage

```
IsingFit(x, family='binomial', AND = TRUE, gamma = 0.25,
plot = TRUE, progressbar = TRUE, lowerbound.lambda = NA,...)
```

Arguments

x	Input matrix. The dimension of the matrix is nobs x nvars; each row is a vector of observations of the variables. Must be cross-sectional data.
family	The default is 'binomial', treating the data as binary. Currently, this procedure is only supported for binary data.
AND	Logical. Can be TRUE of FALSE to indicate whether the AND-rule or the OR-rule should be used to define the edges in the network. Defaults to TRUE.
gamma	A value of hyperparameter gamma in the extended BIC. Can be anything between 0 and 1. Defaults to .25.
plot	Logical. Should the resulting network be plotted?
progressbar	Logical. Should the pbar be plotted in order to see the progress of the estimation procedure?
lowerbound.lambda	The minimum value of tuning parameter lambda (regularization parameter). Can be used to compare networks that are based on different sample sizes. The lowerbound.lambda is based on the number of observations in the smallest group n: sqrt(log(p)/n). p is the number of variables, that should be the same in both groups. When both networks are estimated with the same lowerbound for lambda (based on the smallest group), the two networks can be directly compared.
...	Arguments sent to `qgraph`.

Value

IsingFit returns (invisibly) a 'IsingFit' object that contains the following items:

weiadj	The weighted adjacency matrix.
thresholds	Thresholds of the variables.
q	The object that is returned by qgraph (class 'qgraph').
gamma	The value of hyperparameter gamma.
AND	A logical indicating whether the AND-rule is used or not. If not, the OR-rule is used.
time	The time it took to estimate the network.
asymm.weights	The (asymmetrical) weighted adjacency matrix before applying the AND/OR rule.
lambda.values	The values of the tuning parameter per node that ensured the best fitting set of neighbors.

Note

See also my website: http://cvborkulo.com

Author(s)

Claudia D. van Borkulo, Sacha Epskamp, with contributions from Alexander Robitzsch

Maintainer: Claudia D. van Borkulo <cvborkulo@gmail.com>

References

Chen, J., & Chen, Z. (2008). Extended bayesian information criteria for model selection with large model spaces. Biometrika, 95(3), 759-771.

Foygel, R., & Drton, M. (2011). Bayesian model choice and information criteria in sparse generalized linear models. arXiv preprint arXiv:1112.5635.

Ravikumar, P., Wainwright, M. J., & Lafferty, J. D. (2010). High-dimensional Ising model selection using l1-regularized logistic regression. The Annals of Statistics, 38, 1287 - 1319.

van Borkulo, C. D., Borsboom, D., Epskamp, S., Blanken, T. F., Boschloo, L., Schoevers, R. A., & Waldorp, L. J. (2014). A new method for constructing networks from binary data. Scientific Reports 4, 5918; DOI:10.1038/srep05918.

Examples

```
library("IsingSampler")

### Simulate dataset ###
# Input:
N <- 6 # Number of nodes
nSample <- 1000 # Number of samples

# Ising parameters:
Graph <- matrix(sample(0:1,N^2,TRUE,prob = c(0.8, 0.2)),N,N) * runif(N^2,0.5,2)
Graph <- pmax(Graph,t(Graph))
diag(Graph) <- 0
Thresh <- -rowSums(Graph) / 2

# Simulate:
Data <- IsingSampler(nSample, Graph, Thresh)

### Fit using IsingFit ###
Res <- IsingFit(Data, family='binomial', plot=FALSE)

# Plot results:
library("qgraph")
layout(t(1:2))
qgraph(Res$weiadj,fade = FALSE)
title("Estimated network")
qgraph(Graph,fade = FALSE)
title("Original network")
```

Index