

Package ‘bootnet’

July 27, 2018

Type Package

Title Bootstrap Methods for Various Network Estimation Routines

Version 1.1.0

Author Sacha Epskamp and Eiko I. Fried

Maintainer Sacha Epskamp <mail@sachaepskamp.com>

Depends ggplot2, R (>= 3.0.0)

Imports methods, igraph, IsingFit, qgraph, dplyr (>= 0.3.0.2), tidyr, gtools, corpcor, IsingSampler, mvtnorm, abind, Matrix, parallel, huge, parcor, mgm (>= 1.2), relaimpo, NetworkToolbox (>= 1.1.0), pbapply, graphicalVAR, BDgraph, psych

Description Bootstrap methods to assess accuracy and stability of estimated network structures and centrality indices. Allows for flexible specification of any undirected network estimation procedure in R, and offers default sets for 'qgraph', 'IsingFit', 'IsingSampler', 'glasso', 'huge' and 'parcor' packages.

License GPL-2

NeedsCompilation no

Repository CRAN

Date/Publication 2018-07-27 15:10:02 UTC

R topics documented:

anRpackage-package	2
binarize	2
bootnet	3
corStability	6
differenceTest	7
estimateNetwork	9
genGGM	16
ggmGenerator	17
IsingGenerator	18
netSimulator	18
netSimulator and replicationSimulator methods	22

null	23
plot.bootnet	24
plot.bootnetResult	26
print.bootnet	27
summary.bootnet	27

Index 29

anRpackage-package *Bootstrap Methods for Various Network Estimation Routines*

Description

Bootstrap standard errors on various network estimation routines, such as EBICglasso from the qgraph package and IsingFit from the IsingFit package. See [bootnet](#)

Details

Package: bootnet
 Type: Package
 Version: 0.1
 Date: 2015-03-01
 License: GPL-2

Author(s)

Sacha Epskamp
 Maintainer: Sacha Epskamp <mail@sachaepskamp.com>

See Also

[bootnet](#)

binarize *Binarizes a dataset*

Description

This function will transform data into binary data (0,1). If the data is already binary, this function does nothing.

Usage

```
binarize(x, split = "median", na.rm = TRUE, removeNArows = TRUE, verbose = TRUE)
```

Arguments

x	A data frame or matrix
split	Either a function to split on (as character or as function) or a vector. e.g., split = "mean" will split every variable on the mean of that variable, split=2 will make every value above 2 a 1 and every value below 2 a 0 and a vector of the same length as each variable in the dataset will use those elements to split.
na.rm	The na.rm argument used in the split function.
removeNArows	Logical, should rows with NA be removed?
verbose	Output progress to the console?

Value

A binarized data frame

Author(s)

Sacha Epskamp <mail@sachaepskamp.com>

bootnet

Bootstrapped network estimation

Description

This function can be used to bootstrap network estimation methods so that the spread of parameter and centrality estimates can be assessed. Most important methods are type = 'nonparametric' for the non-parametric bootstrap and type = 'case' for the case-dropping bootstrap. See also Epskamp, Borsboom and Fried (2016) for more details.

Usage

```
bootnet(data, nBoots = 1000, default = c("none", "EBICglasso", "ggmModSelect", "pcor",
    "IsingFit", "IsingSampler", "huge", "adalasso", "mgm", "relimp", "cor", "TMFG",
    "ggmModSelect", "LoGo"), type = c("nonparametric", "parametric", "node", "person",
    "jackknife", "case"), nCores = 1, statistics = c("edge", "strength", "closeness",
    "betweenness", "outStrength", "inStrength"), model = c("detect", "GGM", "Ising",
    "graphicalVAR"), fun, prepFun, prepArgs, estFun, estArgs, graphFun, graphArgs,
    intFun, intArgs, verbose = TRUE, construct = c("default", "function", "arguments"),
    labels, alpha = 1, caseMin = 0.05, caseMax = 0.75, caseN = 10, subNodes, subCases,
    computeCentrality = TRUE, propBoot = 1, replacement = TRUE, graph, sampleSize,
    intercepts, weighted, signed, directed, ...)
```

Arguments

data	A data frame or matrix containing the raw data. Must be numeric, integer or ordered factors.
nBoots	Number of bootstraps
default	A string indicating the method to use. See documentation at estimateNetwork .
type	The kind of bootstrap method to use.
nCores	Number of cores to use in computing results. Set to 1 to not use parallel computing.
statistics	Vector indicating which statistics to store. Can contain "edge", "strength", "closeness", "betweenness", "length", "distance", "expectedInfluence", "inExpectedInfluence", or "outExpectedInfluence". By default, length and distance are not stored.
model	The modeling framework to use. Automatically detects if data is binary or not.
fun	A custom estimation function, when no default set is used. This must be a function that takes the data as input (first argument) and returns either a weights matrix or a list containing the elements "graph" for the weights matrix, "intercepts" for the intercepts (optional) and "results" for the full estimation results (optional).
prepFun	See documentation at estimateNetwork .
prepArgs	See documentation at estimateNetwork .
estFun	See documentation at estimateNetwork .
estArgs	See documentation at estimateNetwork .
graphFun	See documentation at estimateNetwork .
graphArgs	See documentation at estimateNetwork .
intFun	See documentation at estimateNetwork .
intArgs	See documentation at estimateNetwork .
verbose	Logical. Should progress of the function be printed to the console?
labels	A character vector containing the node labels. If omitted the column names of the data are used.
alpha	The centrality tuning parameter as used in centrality .
subNodes	Range of nodes to sample in node-drop bootstrap
caseMin	Minimum proportion of cases to drop when type = "case".
caseMax	Maximum proportion of cases to drop when type = "case".
caseN	Number of sampling levels to test when type = "case".
subCases	Range of persons to sample in person-drop bootstrap
computeCentrality	Logical, should centrality be computed?
propBoot	Proportion of persons to sample in bootstraps. Set to lower than 1 for m out of n bootstrap
replacement	Logical, should replacement be used in bootstrap sampling?

graph	A given network structure to use in parametric bootstrap.
sampleSize	The samplesize to use in parametric bootstrap.
intercepts	Intercepts to use in parametric bootstrap.
construct	How to construct the estimator? Set to "arguments" to use old functionality (not recommended).
weighted	Logical, should the analyzed network be weighted?
signed	Logical, should the analyzed network be signed?
directed	Logical, is the analyzed network directed? Usually does not have to be set and is detected automatically.
...	Additional arguments used in the estimator function.

Value

A bootnet object with the following elements:

sampleTable	A data frame containing all estimated values on the real sample.
bootTable	A data frame containing all estimated values on all bootstrapped samples.
sample	A bootnetResult object with plot and print method containing the estimated network of the real sample.
boots	A list of bootnetResult objects containing the raw bootstrap results.

Author(s)

Sacha Epskamp <mail@sachaepskamp.com>

References

Epskamp, S., Borsboom, D., & Fried, E. I. (2016). Estimating psychological networks and their accuracy: a tutorial paper. arXiv preprint, arXiv:1604.08462.

See Also

[estimateNetwork](#), [differenceTest](#), [corStability](#), [plot.bootnet](#), [summary.bootnet](#)

Examples

```
# BFI Extraversion data from psych package:
library("psych")
data(bfi)
bfiSub <- bfi[,1:25]

# Estimate network:
Network <- estimateNetwork(bfiSub, default = "EBICglasso")

# Centrality indices:
library("qgraph")
centralityPlot(Network)
```

```

## Not run:
# Estimated network:
plot(Network, layout = 'spring')

### Non-parametric bootstrap ###
# Bootstrap 1000 values, using 8 cores:
Results1 <- bootnet(Network, nBoots = 1000, nCores = 8)

# Plot bootstrapped edge CIs:
plot(Results1, labels = FALSE, order = "sample")

# Plot significant differences (alpha = 0.05) of edges:
plot(Results1, "edge", plot = "difference", onlyNonZero = TRUE,
     order = "sample")

# Plot significant differences (alpha = 0.05) of node strength:
plot(Results1, "strength", plot = "difference")

# Test for difference in strength between node "A1" and "C2":
differenceTest(Results, "A1", "C2", "strength")

### Case-drop bootstrap ###
# Bootstrap 1000 values, using 8 cores:
Results2 <- bootnet(Network, nBoots = 1000, nCores = 8,
                   type = "case")

# Plot centrality stability:
plot(Results2)

# Compute CS-coefficients:
corStability(Results2)

## End(Not run)

```

corStability

Correlation stability coefficient

Description

This coefficient denotes the estimated maximum number of cases that can be dropped from the data to retain, with 95% probability, a correlation of at least 0.7 (default) between statistics based on the original network and statistics computed with less cases. This coefficient should not be below 0.25 and is preferably above 0.5. See also Epskamp, Borsboom and Fried (2016) for more details.

Usage

```
corStability(x, cor = 0.7, statistics = c("strength", "closeness", "betweenness"),
            verbose = TRUE)
```

Arguments

x	Output of bootnet . Must be case-drop bootstrap.
cor	The correlation level to test at.
statistics	The statistic(s) to test for.
verbose	Logical, should information on the progress be printed to the console?

Author(s)

Sacha Epskamp <mail@sachaepskamp.com>

References

Epskamp, S., Borsboom, D., & Fried, E. I. (2016). Estimating psychological networks and their accuracy: a tutorial paper. arXiv preprint, arXiv:1604.08462.

See Also

[bootnet](#)

Examples

```
## Not run:
# BFI Extraversion data from psych package:
library("psych")
data(bfi)
bfiSub <- bfi[,1:25]

# Estimate network:
Network <- estimateNetwork(bfiSub, default = "EBICglasso")

# Bootstrap 1000 values, using 8 cores:
# Bootstrap 1000 values, using 8 cores:
Results2 <- bootnet(Network, nBoots = 1000, nCores = 8,
                    type = "case")

# Compute CS-coefficients:
corStability(Results2)

## End(Not run)
```

differenceTest

Bootstrapped difference test

Description

This function computes the bootstrapped difference test for edge-weights and centrality indices. A confidence interval is constructed on the difference of two values, and the test is deemed significant if zero is not in this confidence interval. See also Epskamp, Borsboom and Fried (2016) for more details.

Usage

```
differenceTest(bootobject, x, y, measure = c("strength", "closeness", "betweenness"),
               alpha = 0.05, x2, y2, verbose = TRUE)
```

Arguments

bootobject	Output of bootnet . Must be nonparametric or parametric bootstrap.
x	A character string indicating the ID of a node or an edge, or a number indicating the node or edge. For an edge, can be e.g., "1--2" or "x = 1, x2 = 2".
y	A character string indicating the ID of a node or an edge, or a number indicating the node or edge. For an edge, can be e.g., "1--2" or "y = 1, y2 = 2".
measure	Measure to test. Can be "strength", "closeness", "betweenness", "edge" or "distance".
alpha	Significance level to test at. Note that the actual significance level is influenced by the number of bootstrap samples, and is returned in a message.
x2	Second node in an edge. optional.
y2	Second node in an edge. optional.
verbose	Logical, should the message indicating actual significance level be printed?

Author(s)

Sacha Epskamp <mail@sachaepskamp.com>

References

Epskamp, S., Borsboom, D., & Fried, E. I. (2016). Estimating psychological networks and their accuracy: a tutorial paper. arXiv preprint, arXiv:1604.08462.

See Also

[bootnet](#)

Examples

```
## Not run:
# BFI Extraversion data from psych package:
library("psych")
data(bfi)
bfiSub <- bfi[,1:25]

# Estimate network:
Network <- estimateNetwork(bfiSub, default = "EBICglasso")

# Bootstrap 1000 values, using 8 cores:
Results1 <- bootnet(Network, nBoots = 1000, nCores = 8)

# Test for difference in strength between node "A1" and "C2":
differenceTest(Results, "A1", "C2", "strength")
```



```
# Test for difference between edge N1--N2 and N3--N4:
differenceTest(Results, "N1--N2", "N3--N4", "edge")

# Alternative:
differenceTest(Results, x = "N1", x2 = "N2", y = "N3",
              y2 = "N4", measure = "edge")

## End(Not run)
```

estimateNetwork	<i>Estimate a network structure</i>
-----------------	-------------------------------------

Description

This function allows for flexible estimation of a network structure using various R packages and model frameworks. This is typically done by using one of the default sets. See details for manual specification. See also Epskamp, Borsboom and Fried (2016) for more details.

Usage

```
estimateNetwork(data, default = c("none", "EBICglasso", "pcor", "IsingFit", "IsingSampler",
  "huge", "adalasso", "mgm", "relimp", "cor", "TMFG", "ggmModSelect", "LoGo",
  "graphicalVAR"), fun, prepFun, prepArgs, estFun, estArgs, graphFun, graphArgs,
  intFun, intArgs, labels, verbose = TRUE, construct = c("default", "function",
  "arguments"), .dots = list(), weighted = TRUE, signed = TRUE, directed, datatype,
  checkNumeric = FALSE, ..., .input, memorysaver = FALSE)

bootnet_EBICglasso(data, tuning = 0.5, corMethod = c("cor_auto", "cov", "cor", "nnp"),
  missing = c("pairwise", "listwise", "fiml", "stop"), sampleSize = c("maximum", "minimim"),
  verbose = TRUE, corArgs = list(), refit = FALSE, principalDirection = FALSE,
  lambda.min.ratio = 0.01, nlambda = 100, threshold = FALSE, ...)

bootnet_pcor(data, corMethod = c("cor_auto", "cov", "cor", "nnp"), missing = c("pairwise",
  "listwise", "fiml", "stop"), sampleSize = c("maximum", "minimim"), verbose = TRUE,
  corArgs = list(), threshold = 0, alpha = 0.05, principalDirection = FALSE)

bootnet_cor(data, corMethod = c("cor_auto", "cov", "cor", "nnp"), missing = c("pairwise",
  "listwise", "fiml", "stop"), sampleSize = c("maximum", "minimim"), verbose = TRUE,
  corArgs = list(), threshold = 0, alpha = 0.05, principalDirection = FALSE)

bootnet_IsingFit(data, tuning = 0.25, missing = c("listwise", "stop"), verbose = TRUE,
  rule = c("AND", "OR"), split = "median", principalDirection = FALSE)

bootnet_IsingSampler(data, missing = c("listwise", "stop"), verbose = TRUE,
  split = "median", method = c("default", "ll", "pl", "uni", "bi"),
  principalDirection = FALSE)
```

```

bootnet_adalasso(data,missing = c("listwise","stop"), verbose = TRUE, nFolds = 10,
  principalDirection = FALSE)

bootnet_huge(data, tuning = 0.5, missing = c("listwise","stop"), verbose = TRUE,
  npn = TRUE, criterion = c("ebic","ric","stars"), principalDirection = FALSE,
  lambda.min.ratio = 0.01, nlambda = 100)

bootnet_mgm(data, type, level, tuning = 0.5, missing = c("listwise","stop"),
  verbose = TRUE, criterion = c("EBIC","CV"), nFolds = 10, order = 2,
  rule = c("AND","OR"), binarySign, ...)

bootnet_relimp(data, normalized = TRUE, type = "lmg",
  structureDefault = c("none", "custom", "EBICglasso",
  "pcor", "IsingFit", "IsingSampler", "huge",
  "adalasso", "mgm", "cor", "TMFG", "ggmModSelect",
  "LoGo"), missing = c("listwise", "stop"), ..., verbose
  = TRUE, threshold = 0)

bootnet_TMFG(data, graphType = c("cor", "pcor"), corMethod =
  c("cor_auto", "cov", "cor", "npn"), missing =
  c("pairwise", "listwise", "fiml", "stop"), verbose =
  TRUE, corArgs = list(), principalDirection = FALSE,
  ...)

bootnet_LoGo(data, corMethod = c("cor_auto","cov","cor","npn"), missing = c("pairwise",
  "listwise","fiml","stop"), verbose = TRUE, corArgs = list(), principalDirection = FALSE,
  ...)

bootnet_graphicalVAR(data, tuning = 0.5, verbose = TRUE, principalDirection = FALSE,
  missing = c("listwise","stop"), ...)

bootnet_ggmModSelect(data,tuning = 0, corMethod = c("cor_auto","cov","cor","npn"),
  missing = c("pairwise","listwise","fiml","stop"), sampleSize = c("maximum","minimim"),
  verbose = TRUE, corArgs = list(), principalDirection = FALSE, start = c("glasso",
  "empty","full"), stepwise = TRUE, nCores = 1, ...)

```

Arguments

data	A data frame or matrix containing the raw data. Must be numeric, integer or ordered factors.
default	A string indicating the method to use. Specifying a default sets default values to prepFun, prepArgs, estFun, estArgs, graphFun, graphArgs, intFun and intArgs. Setting a default can be omitted but that does require specifying all above mentioned arguments. Current options are: "EBICglasso" Gaussian Markov random field estimation using graphical LASSO and extended Bayesian information criterion to select optimal regularization parameter. Using EBICglasso from the qgraph package. Calls bootnet_EBICglasso. "IsingFit" Ising model estimation using LASSO regularized nodewise logis-

tic regression and extended Bayesian information criterion to select optimal regularization parameter. Using `IsingFit` from the `IsingFit` package. Calls `bootnet_IsingFit`.

"IsingSampler""}{Calls the `EstimateIsing` function from the `IsingSampler` pack

Uses the `adalasso.net` from the `parcor` package. Calls `bootnet_adalasso`.

"huge" Uses EBIC model selection of GGM networks estimated via the `glasso` algorithm as implemented in the `huge` package (as opposed to `glasso` and `qgraph` packages used in `default = "EBICglasso"`). Uses nonparanormal transformation in preparing the data and does not use polychoric correlations. Calls `bootnet_huge`.

"mgm" Estimates a Mixed graphical model by using the `mgm` (or `mgmfit` in older versions) function of the `mgm` package. Calls `bootnet_mgm`.

"TMFG" Estimates a Triangulated Maximally Filtered Graph, using the function `TMFG` of the `NetworkToolbox` package. Calls `bootnet_TMFG`. Note that this estimates a **correlation network** by default (use the `'graphType'` argument to estimate a partial correlation network instead).

"LoGo" Estimates a Local/Global Sparse Inverse Covariance Matrix, using the function `LoGo` of the `NetworkToolbox` package. Calls `bootnet_LoGo`.

"relimp" Estimates a (directed) relative importance network, using the function `'calc.relimp'` of the `'relaimpo'` package. The `'structureDefault'` argument can be used to use a different default set for estimating the structure of the graph. Calls `bootnet_relimp`.

"ggmModSelect" Estimates an unregularized GGM using the `glasso` algorithm and stepwise model selection, using the `'ggmModSelect'` function from the `qgraph` package. Calls `bootnet_ggmModSelect`.

"graphicalVAR" Estimates a graphical VAR model using the `graphicalVAR` package. This results in two networks which can be plotted using the `'graph'` argument in the `plot` method. Calls `bootnet_graphicalVAR`.

See details section for a more detailed description.

fun	A custom estimation function, when no default set is used. This must be a function that takes the data as input (first argument) and returns either a weights matrix or a list containing the elements "graph" for the weights matrix, "intercepts" for the intercepts (optional) and "results" for the full estimation results (optional).
tuning	EBIC tuning parameter, used in <code>'EBICglasso'</code> , <code>'IsingFit'</code> , <code>'huge'</code> , <code>'mgm'</code> and <code>'ggmModSelect'</code> default sets. Note that the default value differs: <code>'EBICglasso'</code> , <code>'huge'</code> and <code>'mgm'</code> use 0.5, <code>'IsingFit'</code> uses 0.25 and <code>'ggmModSelect'</code> uses 0.
corMethod	Correlation method, used in <code>'EBICglasso'</code> and <code>'pcor'</code> default sets. <code>"cor_auto"</code> uses <code>cor_auto</code> for polychoric and polyserial correlations, <code>"cov"</code> uses the <code>cov</code> function for covariances, <code>"cor"</code> will use the <code>cor</code> function for correlations and <code>"npn"</code> will apply the nonparanormal transformation (via <code>huge.npn</code>) and then compute correlations.
missing	How to handle missing data? <code>"pairwise"</code> for pairwise deletion, <code>"listwise"</code> for listwise deletion, <code>"fiml"</code> for full-information maximum likelihood and <code>"stop"</code> to stop with an error.

sampleSize	How will sample size be computed in EBICglasso default set? "maximum" takes total number of rows including rows with NA, and "minimum" takes total number of rows that contain no NA.
corArgs	A list with arguments for the function used defined by corMethod.
threshold	Thresholding to use in partial correlation networks. Can be a fixed number to threshold all absolute edges below this value, 'locfdr' for local FDR, or any option corresponding to adjustments in <code>corr.p</code> ('none', 'sig', 'holm', 'hochberg', 'hommel', 'bonferroni', 'BH', 'BY' or 'fdr').
refit	Logical used in EBICglasso default set: should the estimated model structure be refitted without LASSO regularization?
rule	The rule to use to select an edge in nodewise estimation. "AND" to only select in edge if both regression coefficients are nonzero and "OR" if only one is nonzero. Used in 'IsingFit' and 'mgm' default sets.
split	A function or character string ("median" or "mean") indicating how to binarize values when estimating an Ising model.
method	The estimation method used in the IsingSampler default set (see EstimateIsing).
nnp	Logical, should nonparanormal be used in huge default set?
criterion	The criterion used in model selection. "ebic", "ric" or "stars" in the huge default set or "EBIC" or "CV" in the mgm default set.
nFolds	Number of folds used in k-fold cross-validation.
type	For mgm, see mgm or mgmfit; for relative importance networks, see <code>calc.relimp</code>
level	See <code>mgm</code> . Automatically set if not assigned.
order	Order up until including which interactions are included in the model. See <code>mgm</code> . Automatically set if not assigned.
binarySign	See <code>mgm</code> . Automatically set if not assigned.
normalized	Should normalized relative importance be used in relative importance networks?
structureDefault	In relative importance networks, default set used to compute the graph structure. Any other arguments used (using ...) are sent to the graph estimator function as well.
graphType	"cor" to estimate a correlation network and "pcor" to estimate a partial correlation network (GGM)
alpha	Significance level to test at.
principalDirection	Rescales variables according to the sign of the first eigen-vector. This will lead to most correlations to be positive (positive manifold), leading to negative edges to be substantively interpretable.
stepwise	Logical indicating if 'ggmModSelect' should use stepwise estimation.
start	See <code>ggmModSelect</code>
prepFun	A function that takes as input the raw data and returns whatever the estimation function needs (as first argument). Typically this function is used to correlate or binarize the data. Defaults to <code>identity</code> if omitted and default is not set. ARGUMENT IS DEPRECATED AND WILL BE REMOVED.

prepArgs	A list with arguments for prepFun. Defaults to list() if omitted and default is not set. ARGUMENT IS DEPRECATED AND WILL BE REMOVED.
estFun	A function that takes as input the prepared raw data and estimates a network. ARGUMENT IS DEPRECATED AND WILL BE REMOVED.
estArgs	A list with arguments for estArgs. ARGUMENT IS DEPRECATED AND WILL BE REMOVED.
graphFun	A function that takes the result of estFun and extracts the estimated weights matrix. Defaults to identity if omitted and default is not set. ARGUMENT IS DEPRECATED AND WILL BE REMOVED.
graphArgs	A list with arguments for graphFun. Defaults to list() if omitted and default is not set. ARGUMENT IS DEPRECATED AND WILL BE REMOVED.
intFun	A function that takes the result of estFun and extracts the estimated intercepts. Defaults to null if omitted and default is not set. ARGUMENT IS DEPRECATED AND WILL BE REMOVED.
intArgs	A list with arguments for intFun. Defaults to list() if omitted and default is not set. ARGUMENT IS DEPRECATED AND WILL BE REMOVED.
labels	A character vector containing the node labels. If omitted the column names of the data are used.
verbose	Logical, currently only used when default = "EBICglasso" in the cor_auto function.
weighted	Logical, should the analyzed network be weighted?
signed	Logical, should the analyzed network be signed?
directed	Logical, is the analyzed network directed? Usually does not have to be set and is detected automatically.
construct	Should the estimator function be constructed using new bootnet functionality (function) or old bootnet functionality (arguments)? Will default to function unless arguments such as estFun are used.
datatype	"normal" if the data argument is a data frame, or "graphicalVAR" if the data argument is a data list that can be used as input to the graphicalVAR package.
checkNumeric	Logical: should the data be checked to be numeric?
lambda.min.ratio	Minimal lambda ratio (LASSO tuning parameter)
nlambda	Number of LASSO tuning parameters to test
nCores	Number of cores to use in estimating networks
.dots	A list of arguments used in the estimation function set by a default set or by the fun argument.
...	A list of arguments used in the estimation function set by a default set or by the fun argument.
.input	Used internally in the bootnet function. Do not use.
memorysaver	Logical. If TRUE attempts to save memory (RAM) by removing some objects from the output. Used by bootnet by default for bootstraps.

Details

NEW FUNCTIONALITY:

The user can manually specify an estimation method by assigning a custom function to the 'fun' argument. This function must take data as input and output an estimated network. The functions `bootnet_` correspond to the functions used when using default sets. E.g, `default = "pcor"` sets `fun = bootnet_pcor`. As such, the arguments described above can be used depending on the default set used.

DEPRECATED OLD FUNCTIONALITY:

The user can manually set the estimation method of the PMRF in `bootnet` using a set of arguments to the `bootnet` function. First, the method of preprocessing the data must be defined via the `prepFun` argument, which must be assigned a function that takes a dataset as input and returns the viable input for the network estimator. The argument `prepArgs` can be specified a list of arguments to the `prepFun` function. Data preprocessing typically means correlating the data for the GGM or binarizing it for the Ising model (to this end `bootnet` provides a `binarize` function). Next, we estimate the network. To do so, we assign the `estFun` argument, a function that takes whatever the output of `prepFun` was and estimates a network. The `estArgs` argument can be used to assign a list of additional arguments to the function used in `estFun`. Finally, we need to extract the weights matrix and intercepts. Assigning functions to the `graphFun` and `intFun` arguments respectively can do this. An example of how these commands work together to estimate a network using `default = "EBICglasso"` is shown below:

```
estimateNetwork(Data,prepFun = cor_auto,prepArgs = list(verbose = FALSE),estFun = qgraph::EBICglasso
```

The following defaults can be used:

```
default = "EBICglasso" prepFun cor_auto from the qgraph package
```

```
  prepArgs list(verbose=TRUE) for estimating a network and list(verbose=FALSE) for
    bootstrap samples.
```

```
  estFun EBICglasso from the qgraph package
```

```
  estArgs list(n = nrow(data), returnAllResults = TRUE)
```

```
  graphFun function(x)x[['optnet']]
```

```
  graphArgs list()
```

```
  intFun null
```

```
  intArgs list()
```

```
default = "IsingFit" prepFun binarize
```

```
  prepArgs list()
```

```
  estFun IsingFit from the IsingFit package
```

```
  estArgs list(plot = FALSE, progress = FALSE)
```

```
  graphFun function(x)x[['weiadj']]
```

```
  graphArgs list()
```

```
  intFun function(x)x[['thresholds']]
```

```
  intArgs list()
```

```
default = "pcor" prepFun cor_auto from the qgraph package
```

```
  prepArgs list()
```

```
  estFun cor2pcor from the corpcor package
```

```
  estArgs list()
```

```

graphFun identity
graphArgs list()
intFun null
intArgs list()
default = "adalasso" prepFun identity
prepArgs list()
estFun adalasso.net from the parcor package
estArgs list()
graphFun function(x)as.matrix(Matrix::forceSymmetric(x$pcor.adalasso))
graphArgs list()
intFun null
intArgs list()
default = "huge" prepFun function(x)huge.npn(na.omit(as.matrix(x)), verbose = FALSE),
  using huge.npn from the huge package
prepArgs list()
estFun function(x)huge::huge.select(huge::huge(x, method = "glasso", verbose=FALSE), criterion = "
  using the huge package
estArgs list()
graphFun as.matrix(qgraph::wi2net(as.matrix(x$opt.icov))), using the huge pack-
  age
graphArgs list()
intFun null
intArgs list()

```

Value

A `bootnetResult` object with the following elements:

<code>graph</code>	The weights matrix of the network
<code>intercepts</code>	The intercepts
<code>results</code>	The results of the estimation procedure
<code>labels</code>	A vector with node labels
<code>nNodes</code>	Number of nodes in the network
<code>nPerson</code>	Number of persons in the network
<code>input</code>	Input used, including the result of the default set used

Author(s)

Sacha Epskamp <mail@sachaepskamp.com>

References

Epskamp, S., Borsboom, D., & Fried, E. I. (2016). Estimating psychological networks and their accuracy: a tutorial paper. arXiv preprint, arXiv:1604.08462.

See Also[bootnet](#)**Examples**

```
# BFI Extraversion data from psych package:
library("psych")
data(bfi)
bfiSub <- bfi[,1:25]

# Estimate network:
Network <- estimateNetwork(bfiSub, default = "EBICglasso")

## Not run:
# Some pointers:
print(Network)

# Estimated network:
plot(Network, layout = 'spring')

# Centrality indices:
library("qgraph")
centralityPlot(Network)

# BIC model selection:
Network_BIC <- estimateNetwork(bfiSub, default = "EBICglasso", tuning = 0)

# Ising model:
Network_BIC <- estimateNetwork(bfiSub, default = "IsingFit")

## End(Not run)
```

genGGM

Generates a GGM small-world network.

Description

Simulates a GGM as described by Yin and Li (2011), using the Watts and Strogatz (1998) algorithm for generating the graph structure (see [watts.strogatz.game](#)).

Usage

```
genGGM(Nvar, p = 0, nei = 1, parRange = c(0.5,1), constant = 1.5, propPositive = 0.5,
       clusters = NULL, graph = c("smallworld", "random", "scalefree", "hub", "cluster"))
```


Arguments

nvar	Number of nodes
p	Rewiring probability if graph = "smallworld" or "cluster", or connection probability if graph = "random". If cluster, can add multiple p's for each cluster, e.g., "c(.1, .5)"
nei	Neighborhood (see watts.strogatz.game).
parRange	Range of partial correlation coefficients to be originally sampled.
constant	A constant as described by Yin and Li (2011).
propPositive	Proportion of edges to be set positive.
clusters	Number of clusters if graph = "cluster"
graph	Type of graph to simulate

Author(s)

Sacha Epskamp <mail@sachaepskamp.com>

References

- Yin, J., and Li, H. (2011). A sparse conditional gaussian graphical model for analysis of genetical genomics data. *The annals of applied statistics*, 5(4), 2630.
- Watts, D. J., & Strogatz, S. H. (1998). Collective dynamics of 'small-world' networks. *nature*, 393(6684), 440-442.

ggmGenerator	<i>Generates a function that simulates data from the Gaussian graphical model (GGM)</i>
--------------	---

Description

Generates data given a partial correlation network. Data can be made ordinal by using a threshold model with equally spaced thresholds.

Usage

```
ggmGenerator(ordinal = FALSE, nLevels = 4)
```

Arguments

ordinal	Logical, should ordinal data be generated?
nLevels	Number of levels used in ordinal data.

Author(s)

Sacha Epskamp <mail@sachaepskamp.com>

IsingGenerator *Generates a function that simulates data from the Ising model*

Description

Uses [IsingSampler](#) to generate the data.

Usage

```
IsingGenerator(...)
```

Arguments

... Arguments passed to [IsingSampler](#)

Value

A function with as first argument the sample size and as second argument a named list, with element graph encoding a weights matrix and element intercepts encoding thresholds.

Author(s)

Sacha Epskamp <mail@sachaepskamp.com>

netSimulator *Network Estimation Performance*

Description

This function can be used to run a simulation study on the performance of network estimation by varying sample size or any argument used as input to [estimateNetwork](#). The purpose of this function is to provide a way to assess the required sample size given a network structure, as well as to easily perform simulation studies. By default, the function uses [genGGM](#) to simulate a chain graph or small-world network. See details for more information. The [replicationSimulator](#) function instead assesses how well a network based on a second independent sample would replicate the network based on the first independent sample.

Usage

```
netSimulator(  
  input = genGGM(Nvar = 10),  
  nCases = c(50, 100, 250, 500, 1000, 2500),  
  nReps = 100,  
  nCores = 1,  
  default,  
  dataGenerator,
```

```

    ...,
    moreArgs = list(),
    moreOutput = list()

replicationSimulator(
  input = genGGM(Nvar = 10),
  nCases = c(50, 100, 250, 500, 1000, 2500),
  nReps = 100,
  nCores = 1,
  default,
  dataGenerator,
  ...,
  moreArgs = list())

```

Arguments

input	Either a weights matrix, a list containing elements graph (encoding the weights matrix) and intercepts (encoding the intercepts), or a function generating such objects. By default, <code>genGGM</code> is used to generate a Gaussian graphical model. However, it is recommended to replace this with a prior expected graph structure.
nCases	The sample sizes to test for.
nReps	Number of repetitions per sampling level.
nCores	Number of cores to use. Set to more than 1 to use parallel computing.
default	Default set used (see <code>estimateNetwork</code>). In most cases, this will set <code>dataGenerator</code> to the relevant generator.
dataGenerator	A function that generates data. The first argument must be the sample size, the second argument must be the output of <code>input</code> . Can often be ignored if <code>default</code> is set.
moreArgs	A named list of arguments to be used when estimating the network, but which should not be interpreted as different conditions. Use this argument to assign arguments that require vectors.
moreOutput	List with functions that take the estimated weights matrix as first argument and the true weights matrix as second argument to produce some output.
...	Arguments used by <code>estimateNetwork</code> to estimate the network structure. Providing a vector for any argument will simulate under each value. This way, any argument in <code>estimateNetwork</code> can be used in a simulation study.

Details

any argument to `estimateNetwork` can be used in a simulation study, with a vector (e.g., `rule = c("AND", "OR")`) specifying that both conditions are tested. Adding too many conditions can quickly make any simulation study intractable, so only vary some arguments! The `dataGenerator` argument can be any function that generates data. Currently, only `ggmGenerator` and `IsingGenerator` are implemented in `bootnet` itself, which generates data given a Gaussian graphical model.

Author(s)

Sacha Epskamp <mail@sachaepskamp.com>

Examples

```
# 5-node GGM chain graph:
trueNetwork <- genGGM(5)

# Simulate:
Res <- netSimulator(trueNetwork, nReps = 10)

# Results:
Res

# Plot:
plot(Res)

## Not run:
library("bootnet")

# BFI example:
# Load data:
library("psych")
data(bfi)
bfiData <- bfi[,1:25]

# Estimate a network structure, with parameters refitted without LASSO regularization:
library("qgraph")
Network <- EBICglasso(cor_auto(bfiData), nrow(bfiData), refit = TRUE)

# Simulate 100 repetitions in 8 cores under different sampling levels:
Sim1 <- netSimulator(Network,
                    default = "EBICglasso",
                    nCases = c(100,250,500),
                    nReps = 100,
                    nCores = 8)

# Table of results:
Sim1

# Plot results:
plot(Sim1)

# Compare different default set at two sampling levels:
Sim2 <- netSimulator(Network,
                    default = c("EBICglasso","pcor","huge"),
                    nCases = c(100,250,500),
                    nReps = 100,
                    nCores = 8)

# Print results:
Sim2

# Plot results:
plot(Sim2, xfacet = "default", yvar = "correlation")
```

```
# Difference using polychoric or pearson correlations in ordinal data:
Sim3 <- netSimulator(Network,
  dataGenerator = ggmGenerator(ordinal = TRUE, nLevels = 4),
  default = "EBICglasso",
  corMethod = c("cor", "cor_auto"),
  nCases = c(100, 250, 500),
  nReps = 100,
  nCores = 8)

# Print results:
Sim3

# Plot results:
plot(Sim3, color = "corMethod")

# Ising model:
trueNetwork <- read.csv('http://sachaepskamp.com/files/weiadj.csv')[,-1]
trueNetwork <- as.matrix(trueNetwork)
Symptoms <- rownames(trueNetwork) <- colnames(trueNetwork)
Thresholds <- read.csv('http://sachaepskamp.com/files/thr.csv')[,-1]

# Create an input list (intercepts now needed)
input <- list(graph=trueNetwork, intercepts=Thresholds)

# Simulate under different sampling levels:
Sim4 <- netSimulator(
  input = input,
  default = "IsingFit",
  nCases = c(250, 500, 1000),
  nReps = 100,
  nCores = 8)

# Results:
Sim4

# Plot:
plot(Sim4)

# Compare AND and OR rule:
Sim5 <- netSimulator(
  input = input,
  default = "IsingFit",
  nCases = c(250, 500, 1000),
  rule = c("AND", "OR"),
  nReps = 100,
  nCores = 8)

# Print:
Sim5

# Plot:
plot(Sim5, yfacet = "rule")
```

```
## End(Not run)
```

```
netSimulator and replicationSimulator methods
      netSimulator S3 methods
```

Description

Plot, print and summary methods for `netSimulator` output.

Usage

```
## S3 method for class 'netSimulator'
plot(x, xvar = "factor(nCases)", yvar = c("sensitivity",
      "specificity", "correlation"), xfacet = "measure",
      yfacet = ".", color = NULL, ylim = c(0, 1), print =
      TRUE, xlab = "Number of cases", ylab, outlier.size =
      0.5, boxplot.lwd = 0.5, style = c("fancy", "basic"),
      ...)
## S3 method for class 'netSimulator'
print(x, digits = 2, ...)
## S3 method for class 'netSimulator'
summary(object, digits = 2, ...)

## S3 method for class 'replicationSimulator'
plot(x, yvar = c("correlation", "jaccard",
      "replicatedEdges", "replicatedZeroes"), ...)
## S3 method for class 'replicationSimulator'
print(x, digits = 2, ...)
## S3 method for class 'replicationSimulator'
summary(object, digits = 2, ...)
```

Arguments

<code>x</code>	Output of <code>netSimulator</code> .
<code>object</code>	Output of <code>netSimulator</code> .
<code>xvar</code>	String indicating the variable to be used on the x-axis.
<code>yvar</code>	String vector indicating the variable(s) to be used on the y-axis.
<code>xfacet</code>	String indicating the variable to be used on the horizontal facets (or "." to omit).
<code>yfacet</code>	String indicating the variable to be used on the vertical facets (or "." to omit).
<code>color</code>	String indicating the variable to be used in coloring boxplots.
<code>ylim</code>	Y-axis limits.
<code>print</code>	Logical, should the plot be printed? This helps printing the plots to PDF files.
<code>digits</code>	Number of digits to be used in print and summary method.

xlab	X-axis label
ylab	Y-axis label. Defaults to combining the values in yvar. Is hidden when xfacet = "measure", as then it is clear what the y-axis represent from the facet labels.
outlier.size	Size of the outliers as plotted in boxplots.
boxplot.lwd	Line width of the boxplots
style	"fance" for a style including several aesthetic enhancements, and "basic" for a as simple as possible style.
...	Arguments sent to "plot.netSimulator" from "plot.replicationSimulator"

Author(s)

Sacha Epskamp <mail@sachaepskamp.com>

null

Returns NULL

Description

This function simply returns NULL.

Usage

```
null(...)
```

Arguments

... Anything

Author(s)

Sacha Epskamp <mail@sachaepskamp.com>

Examples

```
null("Not NULL")
```

plot.bootnet

*Plots bootnet results***Description**

This function can be used to plot bootnet results by plotting all bootstrapped statistics as line or by plotting confidence intervals.

Usage

```
## S3 method for class 'bootnet'
plot(x, statistics, plot, graph, CIstyle = c(
  "quantiles", "SE"), rank = FALSE, sampleColor =
  "darkred", samplelwd = 1, meanColor = "black", meanlwd
= 0.5, bootColor = "black", bootAlpha = 0.01, bootlwd
= 0.9, areaAlpha = 0.1, order = c("id", "sample",
"mean"), decreasing = TRUE, perNode = FALSE,
legendNcol = 2, labels = TRUE, legend = TRUE,
subsetRange = c(100, 0), area = !perNode, alpha =
0.05, onlyNonZero = FALSE, differenceShowValue,
differenceEdgeColor = TRUE, verbose = TRUE, panels =
TRUE, split0 = FALSE, prop0 = ifelse(split0, TRUE,
FALSE), prop0_cex = 1, prop0_alpha = 0.8,
prop0_minAlpha = 0.25, subset, ...)
```

Arguments

x	A bootnet object
statistics	The types of statistics to plot. c("intercept", "strength", "closeness", "betweenness") plots all nodewise statistics whereas c("edge", "distance") plots all pairwise statistics. Defaults to "edge" for regular bootstrap and c("strength", "closeness", "betweenness") for node and person drop bootstrap.
plot	Character string indicating what to plot. Can be "area" to produce a graph with the area indicating the confidence region, or "difference" producing a plot showing significant differences. Other options are "line" and "interval", which are currently unstable and not recommended to use.
graph	If multiple graphs are estimated, which graph should be plotted? Currently used for default = "graphicalVAR" to plot a temporal network using graph = "temporal" or a contemporaneous network using graph = "contemporaneous"
CIstyle	Style of CIs to construct. "SE" shows the sample statistic plus and minus two times the standard deviation of bootstraps, and "quantiles" the area between the 2.5th and 97.5th quantile. Defaults to "quantiles".
rank	Logical, should plots show rank of statistics instead of statistics?
sampleColor	Color of the original sample line
samplelwd	Line width of the original sample line

bootColor	Color of the bootstrap lines
bootAlpha	Alpha of the bootstrap lines
bootlwd	Line width of the bootstrap lines
areaAlpha	Alpha of the area
order	String indicating how to order nodes. "id" will order nodes based on their name, "mean" will order nodes based on the average bootstrapped value of the first statistic in statistics, and "sample" will order the nodes as done in "mean" but orders ties based on their sample value.
decreasing	Logical indicating if the ordering is decreasing or increasing.
perNode	Logical, should centrality estimates per node be plotted instead of correlation with original parameter. Only used in node and person drop bootstrap.
legendNcol	Number of columns in the legend if perNode = TRUE.
labels	Logical, should labels be plotted?
legend	Logical, should the legend be plotted?
subsetRange	Range in percentages of the x-axis in node and person drop plots.
area	Logical, should the confidence area be plotted?
alpha	Significance level used in plot = "difference".
onlyNonZero	Logical used when plot = "difference", statistics = "edge", should only edges be included that were nonzero in the estimated network structure?
differenceShowValue	Logical used when plot = "difference". Should values be shown in the diagonal of the difference plot?
differenceEdgeColor	Logical used when plot = "difference", statistics = "edge". Should diagonal blocks be colored according to default edge colors?
verbose	Should expected alpha be printed?
panels	Logical, should panel titles be printed?
meanColor	Color of the bootstrap means.
meanlwd	Line width of the bootstrap means
split0	Logical. When set to TRUE, the displayed intervals are based on occasions when the parameter was not estimated to be zero, and an extra box is added indicating the number of times a parameter is estimated to be zero.
prop0	Logical, should boxes indicating the proportion of times parameters were estimated to be zero be added to the plot?
prop0_cex	Size of the boxes indicating number of times a parameter was set to zero.
prop0_alpha	Transparency of the boxes indicating number of times a parameter was set to zero.
prop0_minAlpha	Minimal transparency of the *lines* of plotted intervals as the proportion of times an edge was not included goes to 0.
subset	Vector indicating labels of nodes to include in the plot. This can be used to show, for example, only edges related to one particular node.
...	Not used.

Value

A ggplot2 object.

Author(s)

Sacha Epskamp <mail@sachaepskamp.com>

plot.bootnetResult *Plot method for bootnetResult objects*

Description

Plots the graph using the qgraph package and the [qgraph](#) function. Defined as `qgraph::qgraph(x[['graph']], labels=x[[`

Usage

```
## S3 method for class 'bootnetResult'
plot(x, graph, weighted, signed, directed, labels,
      layout = "spring", parallelEdge = TRUE, cut = 0, theme = "colorblind", ...)
```

Arguments

x	A bootnetResult object
graph	Numeric or string indicating which graph to plot. Only needed when multiple graphs are estimated. For example, when using <code>default = "graphicalVAR"</code> , <code>graph = "temporal"</code> plots the temporal network and <code>graph = "contemporaneous"</code> plots the contemporaneous network.
weighted	Logical, should the analyzed network be weighted?
signed	Logical, should the analyzed network be signed?
directed	Logical, is the analyzed network directed? Usually does not have to be set and is detected automatically.
labels	Labels of the nodes. Defaults to the column names of the data if missing.
layout	Placement of the nodes. See qgraph . Always defaults to "spring".
parallelEdge	Should edges in directed networks be plotted parallel? See qgraph . Defaults to TRUE instead of FALSE (as in qgraph).
cut	Should scaling in width and saturation of edges be split? See qgraph . Defaults to 0 to disable cut (qgraph chooses a cutoff with at least 20 nodes).
theme	Theme of the edge and node colors. See qgraph . Defaults to "colorblind" rather than the default used in qgraph ("classic").
...	Arguments sent to qgraph .

Author(s)

Sacha Epskamp <mail@sachaepskamp.com>

print.bootnet	<i>Print method for bootnet and bootnetResult objects</i>
---------------	---

Description

Prints a short overview of the results of [bootnet](#)

Usage

```
## S3 method for class 'bootnet'  
print(x, ...)  
## S3 method for class 'bootnetResult'  
print(x, ...)  
## S3 method for class 'bootnetResult'  
summary(object, ...)
```

Arguments

x	A bootnet or bootnetResult object
object	A bootnetResult object
...	Not used.

Author(s)

Sacha Epskamp <mail@sachaepskamp.com>

summary.bootnet	<i>Summarize bootnet results</i>
-----------------	----------------------------------

Description

Creates a data frame (wrapped as [tbl_df](#)) containing summarized results of the bootstraps.

Usage

```
## S3 method for class 'bootnet'  
summary(object, graph, statistics = c("edge", "intercept",  
                                     "strength", "closeness", "betweenness", "distance"),  
        perNode = FALSE, rank = FALSE, tol =  
        sqrt(.Machine$double.eps), ...)
```

Arguments

object	A bootnet object
graph	Numeric or string indicating which graph to summarize. Only needed when multiple graphs are estimated. For example, when using default = "graphicalVAR", graph = "temporal" plots the temporal network and graph = "contemporaneous" plots the contemporaneous network.
statistics	The types of statistics to include in the summary table
perNode	Logical, should centrality estimates per node be plotted instead of correlation with original parameter. Only used in node and person drop bootstrap.
rank	Logical, should plots show rank of statistics instead of statistics?
tol	Tolerance level for assuming an edge is set to zero.
...	Not used.

Value

A tbl_df (data frame) containing summarized statistics.

Author(s)

Sacha Epskamp <mail@sachaepskamp.com>

Index

adalasso.net, [11](#), [15](#)
anRpackage (anRpackage-package), [2](#)
anRpackage-package, [2](#)

binarize, [2](#), [14](#)
bootnet, [2](#), [3](#), [7](#), [8](#), [16](#), [27](#)
bootnet_adalasso (estimateNetwork), [9](#)
bootnet_cor (estimateNetwork), [9](#)
bootnet_EBICglasso (estimateNetwork), [9](#)
bootnet_ggmModSelect (estimateNetwork), [9](#)
bootnet_graphicalVAR (estimateNetwork), [9](#)
bootnet_huge (estimateNetwork), [9](#)
bootnet_IsingFit (estimateNetwork), [9](#)
bootnet_IsingSampler (estimateNetwork), [9](#)
bootnet_LoGo (estimateNetwork), [9](#)
bootnet_mgm (estimateNetwork), [9](#)
bootnet_pcor (estimateNetwork), [9](#)
bootnet_relimp (estimateNetwork), [9](#)
bootnet_TMFG (estimateNetwork), [9](#)

calc.relimp, [12](#)
centrality, [4](#)
cor, [11](#)
cor2pcor, [11](#), [14](#)
cor_auto, [11](#), [14](#)
corr.p, [12](#)
corStability, [5](#), [6](#)
cov, [11](#)

differenceTest, [5](#), [7](#)

EBICglasso, [10](#), [14](#)
EstimateIsing, [11](#), [12](#)
estimateNetwork, [4](#), [5](#), [9](#), [18](#), [19](#)

genGGM, [16](#), [18](#), [19](#)
ggmGenerator, [17](#), [19](#)
ggmModSelect, [12](#)

huge.npn, [11](#), [15](#)

identity, [12](#), [13](#), [15](#)
IsingFit, [11](#), [14](#)
IsingGenerator, [18](#), [19](#)
IsingSampler, [18](#)

mgm, [12](#)

netSimulator, [18](#), [22](#)
netSimulator and replicationSimulator
methods, [22](#)
null, [13](#), [14](#), [23](#)

plot.bootnet, [5](#), [24](#)
plot.bootnetResult, [26](#)
plot.netSimulator (netSimulator and
replicationSimulator methods),
[22](#)
plot.replicationSimulator
(netSimulator and
replicationSimulator methods),
[22](#)
print.bootnet, [27](#)
print.bootnetResult (print.bootnet), [27](#)
print.netSimulator (netSimulator and
replicationSimulator methods),
[22](#)
print.replicationSimulator
(netSimulator and
replicationSimulator methods),
[22](#)

qgraph, [26](#)

replicationSimulator (netSimulator), [18](#)

summary.bootnet, [5](#), [27](#)
summary.bootnetResult (print.bootnet),
[27](#)

summary.netSimulator (netSimulator and
replicationSimulator methods),
[22](#)

summary.replicationSimulator
(netSimulator and
replicationSimulator methods),
[22](#)

tbl_df, [27](#)

watts.strogatz.game, [16](#), [17](#)