Package ‘c3net’

February 19, 2015

Version 1.1.1
Title Infering large-scale gene networks with C3NET
Date 2011-01
Author Gokmen Altay, Frank Emmert-Streib
Description This package allows inferring gene regulatory networks with direct physical interactions from microarray expression data using C3NET.
Maintainer Gokmen Altay <altayscience@gmail.com>
License GPL (>= 3)
Depends R (>= 2.12.1), igraph
Repository CRAN
Repository/R-Forge/Project c3net
Repository/R-Forge/Revision 23
Date/Publication 2012-07-23 10:35:32
NeedsCompilation no

R topics documented:

<table>
<thead>
<tr>
<th>Topic</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>c3</td>
<td>2</td>
</tr>
<tr>
<td>c3net</td>
<td>3</td>
</tr>
<tr>
<td>checknet</td>
<td>4</td>
</tr>
<tr>
<td>copula</td>
<td>6</td>
</tr>
<tr>
<td>expdata</td>
<td>7</td>
</tr>
<tr>
<td>makemim</td>
<td>8</td>
</tr>
<tr>
<td>netplot</td>
<td>9</td>
</tr>
<tr>
<td>sigtestMTC</td>
<td>10</td>
</tr>
<tr>
<td>sigtestp</td>
<td>11</td>
</tr>
<tr>
<td>truenet</td>
<td>12</td>
</tr>
</tbody>
</table>

Index 14
c3

**C3NET second step: Selecting maximum valued elements**

**Description**

c3 takes the mutual information matrix as input and implements the second step of C3NET - see details.

**Usage**

c3( mim, sym = TRUE )

**Arguments**

- **mim**
  - A symmetric square mutual information matrix, where the elements (i,j) correspond to the mutual information $I(i, j)$ or $I(j, i)$ between variables i and j.

- **sym**
  - Decides the output matrix to be symmetric or not.

**Details**

The C3NET algorithm consists of two main steps. The first step is the same as for relevance networks (RELNET), where all the non-significant mutual information values in the matrix are eliminated if statistically not significant. The second step of C3NET keeps all maximum valued mutual information values for each row in the matrix and sets the rest of the elements in the matrix zero (the diagonal of the matrix is ignored). The output is normally symmetric matrix but if the argument sym is set to FALSE then the output becomes non-symmetric.

**Value**

c3 implements the second step of the C3NET algorithm returning a symmetric mutual information matrix. Specifically, the non-zero elements in the returned matrix represent undirected links between variables which are statistically significant (tested in the first step of C3NET).

**References**


**See Also**

makemim, copula, sigtestp, sigtestMTC, netplot, c3net
Examples

data(expdata)

expdata <- copula(expdata)
mim <- makemim(expdata)

Ic <- mean(mim[upper.tri(mim)])  #Example cut-off for the first step of C3NET

# Ic <- 2 can be set for the example.
mim[mim < Ic] <- 0  #nonsignificant values eliminated wrt C3NET step 1.

net <- c3(mim)  # regulatory network inferred (non zero elements stand for links of #the predicted network)

---

c3net  
All in one function to infer network with C3NET

Description

c3net takes a data set as input and computes the inferred network using C3NET - see details.

Usage

c3net(dataset, cop=TRUE, alpha=0.01, methodstep1="cutoff", cutoffMI= 0,
MTCmethod="BH", itnum=5, network=FALSE)

Arguments

dataset  
Data set where rows are variables (e.g. genes) and columns are samples.

cop  
The data set is copula transformed by default. If not wanted, then "cop" is set to FALSE.

alpha  
Statistical significance threshold

itnum  
Number of iterations to resample data to get sampling distribution.

methodstep1  
If it is "cutoff" then the input cutoffMI is used as threshold for elimination in step 1. If it is "MTC" then MTCmethod input is used as the specific multiple testing correction method to employ. If it is "justp" then only significance threshold alpha is used to compute significance threshold MI.

MTCmethod  
Available option are "holm", "hochberg", "hommel", "bonferroni", "BH", "BY".

cutoffMI  
If methodstep1 is set to "cutoff" then this value is used for MI threshold. If this is set to 0 (or by default) it uses mean MI for threshold.

network  
If TRUE, the network is plotted in the end.
Details

For Step 1 of C3NET, there are three options that can be used. If methodstep1="cutoff" then cutoffMI input is taken as the significance threshold for step 1. In case cutoffMI is set to 0 then by default mean MI is taken as cutoffMI. If methodstep1 is set to "MTC" then the method from input MTCmethod, itnum for iteration number to get null distribution and alpha for statistical significance is taken for computations. If methodstep1 is set to "justp" or anything else then just alpha and itnum is used for computations.

Value

c3net returns a symmetric mutual information matrix, which is obtained after implementing C3NET. Specifically, non-zero elements in the returned matrix represents undirected link between variables. The inferred network may also be plotted if the argument network is set TRUE.

References


See Also

makemim, copula.c3, sigtestp, sigtestMTC

Examples

data(expdata)
data(truenet)

net <- c3net(expdata, network=TRUE)
scores <- checknet(net, truenet)

checknet  Validating the inferred network

Description

checknet takes the inferred mutual information matrix and the true network for validation. - see details.

Usage

checknet( finalrelationmatrix, realrelationmatrix)
checknet

Arguments

finalrelationmatrix
The inferred symmetric mutual information matrix, where i,j th element is the mutual information \( I(i, j) \) or \( I(j, i) \) between two variables i and j. The diagonal is set to zero.

realrelationmatrix
A symmetric reference connection matrix, where a 1 at i,j th element defines the connection between variables i and j and non-connection is represented by 0. The diagonal is all zero.

Value

checknet returns a vector with 6 elements, that contains the parameters as output <- c(precision, Fscore, recall, TP, FP, FN).

References


See Also

makemim, copula, c3, c3net, c3, makemim, copula, sigtestp, sigtestMTC

Examples

data(expdata)
data(truenet)
expdata <- copula(expdata)
mim <- makemim(expdata)
Ic <- 2  #Example cut-off value for mutual information for the first step of C3NET
mim[mim < Ic] <- 0  #nonsignificant values eliminated wrt C3NET step 1.
net <- c3(mim)  # regulatory network inferred (non zero elements stand for links of # the predicted network)
scores <- checknet(net, truenet)
copula  Copula transformation of data set

Description

copula takes a data set as input and applies a copula transformation - see details.

Usage

copula( expdata )

Arguments

expdata  Data set; rows correspond to variables (e.g. genes) and columns to samples.

Details

The data set can be copula transformed for more stable estimations of the mutual information matrix. Each row of the data set (gene samples of each gene) is copula transformed and obtained rdata. Here, the formula for conversion is copuladata = (rdata - 0.5)/numberofsamples.

Value

copula returns a copula transformed data set.

References


See Also

c3, c3net, makemim

Examples

data(expdata)

expdata <- copula(expdata)
expdata

Example data set

Description
expdata This data set has been generated by SynTRen. The source network corresponds to a subnetwork of E.coli (file name "EColi_full_HongWu_Ma_NAR2004.sif" from Ma HW, et. al. Nuc. Ac. Res. 2004).

Usage
data(expdata)

Arguments
expdata Example data; rows correspond to variables (e.g. genes) and columns to samples.

Details
The example is also given to provide a sample data set so that one can take it as reference while formatting the new data set as entry to c3net. Data sets must be complete and no empty or non-numerical character allowed.

Value
expdata returns an example data set where rows are variables (e.g. genes) and columns are samples. It consists of 400 genes, 898 edges, and 800 steady-state samples.

References

See Also
c3, c3net, makemim

Examples
data(expdata)
Description

`makemim` takes the data set as input and computes mutual information values for each pair of variables (e.g. gene). - see details.

Usage

`makemim(expdata)`

Arguments

- `expdata`: Data matrix where rows correspond to variables (e.g. genes) and columns to samples.

Details

A mutual information matrix is generated from a data set using an empirical Gaussian estimator.

Value

`makemim` returns a symmetric mutual information matrix using empirical Gaussian estimator.

References


See Also

c3,c3net, copula, checknet

Examples

data(expdata)

expdata <- copula(expdata)

mim <- makemim(expdata)
netplot

Plotting the inferred network

Description

netplot takes the inferred mutual information matrix and plots a network wrt nonzero elements of
the matrix - see details.

Usage

netplot(gnet)

Arguments

gnet  A symmetric squared mutual information matrix that is inferred as final con-
        nection matrix, where i,j th element is the mutual information $I(i, j)$ or $I(j, i)$
        between variables i and j. The nonzero elements may also be 1. Diagonal is all
        zero.

Details

A labelled undirected network is plotted wrt the row (column) names of the input matrix. When
plotted, select "Layout" and then "Fruchterman-Reingold" and "OK" to get a better view. To remove
labels, select "View" and deselect "Labels". To change the colors or size of nodes and edges, from
"Select", click on "all edges" or "vertices" and on the network right click on nodes and change color
or size.

Value

netplot returns an undirected network.

References

G. Altay, F. Emmert-Streib, "Inferring the conservative causal core of gene regulatory networks",

See Also

makemim, copula, c3, sigtestp, sigtestMTC

Examples

data(expdata)

data(truenet)

expdata <- copula(expdata)
mim <- makeMim(expdata)

Ic <- 2  # Example cut-off for the first step of C3NET

mim[mim < Ic] <- 0  # Nonsignificant values eliminated wrt C3NET step 1.

net <- c3(mim)  # Regulatory network inferred (non zero elements stand for links of
# the predicted network)

netplot(net)

---

**Description**

`sigtestMTC` takes a data set as input and computes the mutual information matrix in which the
nonsignificant elements are eliminated by a multiple testing correction - see details.

**Usage**

```r
sigtestMTC(data, alpha, itnum, methodsig="BH")
```

**Arguments**

- **data**: Data set where rows correspond to variables (e.g. genes) and columns to samples.
- **alpha**: Significance level
- **itnum**: Number of iterations to resample data to estimate sampling distribution.
- **methodsig**: A procedure to perform a multiple testing correction, either controlling the false
discovery rate (FDR) or the family-wise error. Available option are "holm",
"hochberg", "hommel", "bonferroni", "BH", "BY".

**Details**

The data set is resampled for each iteration and at the end the sampling distribution is obtained.

**Value**

`sigtestMTC` returns an environment `res` that contains the new MI matrix, `res$Inew`, that is obtained
after the elimination of nonsignificant elements with a MTC procedure. For advanced users, who
are able to write code in R, we provided the variables obtained in the function so that one might
want to make further analysis without running the function again. Please see the source code for
the additional return values of the environment `res`. 

---

**Notes**

The `sigtestMTC` function is designed to help in the analysis of large data sets by applying a multiple
testing correction to the resulting mutual information matrix. This is particularly useful in the context of
network inference, where the identification of significant edges is crucial for understanding biological
processes.
sigtestp

References

See Also
makemim, copula.c3, sigtestp

Examples
data(expdata)
data(truenet)
alpha <- 0.001
itnum <- 2
res <- sigtestMTC( expdata, alpha, itnum, methodsig="BH")
net <- c3(res$Inew) # regulatory network inferred (non zero elements stand for links of # the predicted network)
scores <- checknet(net, truenet)

---
sigtestp  Significance test for elimination of nonsignificant edges

Description
sigtestp takes the data set as input and computes the mutual information matrix in which the nonsignificant elements are eliminated by using a significance threshold only (no MTC) - see details.

Usage
sigtestp( data, alpha, itnum)

Arguments
data  Data set where rows are variables (e.g. genes) and columns are samples.
alphabetical   Statistical significance threshold
itnum         Number of iterations to resample data to get sampling distribution.

Details
The data set is resampled completely at each iteration and at the end the sampling distribution is obtained for using in significance test.
Value

`sigtestp` returns environment `res` that contains the obtained threshold value `res$I0`. For advanced users, who are able to write code in R, we provided the variables obtained in the function so that one might want to make further analysis without running the function again. Please see the source code for the additional return values of the environment `res`.

References


See Also

`makemim`, `copula.c3`, `cS`, `cS`, `sigtestMTC`

Examples

```r
data(expdata)
data(truenet)
alpha <- 0.001
itnum <- 2
res <- sigtestp( expdata, alpha, itnum)
net <- c3(res$Inew) # regulatory network inferred (non zero elements stand for links of the predicted network)
scores <- checknet(net, truenet)
```

---

**truenet**

*Reference, e.g. true, network of the example data set*

Description

`truenet` A symmetric matrix, where non-zero elements correspond to true connections.

Usage

```r
data( truenet )
```

Arguments

truenet Reference (true) network of the example data set.
Details

A symmetric matrix, where non-zero elements correspond to true connections.

Value

`truenet` Reference (true) network of the example data set.

References


See Also

`c3, c3net, makemim`

Examples

`data(truenet)`
Index

*Topic misc
  c3, 2
  c3net, 3
  checknet, 4
  copula, 6
  expdata, 7
  makemim, 8
  netplot, 9
  sigtestMTC, 10
  sigtestp, 11
  truenet, 12

c3, 2, 4–9, 11–13
  c3net, 2, 3, 5–8, 13
  checknet, 4, 8
  copula, 2, 4, 5, 6, 8, 9, 11, 12

expdata, 7

makemim, 2, 4–7, 8, 9, 11–13

netplot, 2, 9

sigtestMTC, 2, 4, 5, 9, 10, 12
  sigtestp, 2, 4, 5, 9, 11, 11

truenet, 12