Package 'lem'

August 6, 2018

Title Linear Effects Models

Version 1.0

Author Ewa Szczurek

Description Methods for simulating and inferring linear effects models

Reference Szczurek E, Beerenwinkel N, Linear effects models of signaling pathways from combinatorial perturbation data, Bioinformatics. 2016 Jun 15;32(12):i297-i305. doi: 10.1093/bioinformatics/btw268.

Maintainer Ewa Szczurek <szczurek@mimuw.edu.pl>

Depends R (>= 3.0)

Imports nem, methods, stats

biocViews Bioinformatics, GraphsAndNetworks, Pathways, SystemsBiology, NetworkInference

URL http://www.bioconductor.org

License GPL (>= 2)

R topics documented:

`lem` *Linear Effects Models - main function*

Description

The main function to perform model estimation from data

Usage

```
lem(Y, E, Gs=NULL, inference="greedy", parameter.estimation="linear.reg", verbose=FALSE)
```

Arguments

Y	Effect measurements: can be a vector (for one measured phenotype) or a matrix (the different phenotypes then stand for columns). Contains continuous values. Each element of the vector or each row of the matrix corresponds to a perturbation experiment under which the measurement was taken.
E	Binary experiment matrix. Each row corresponds to an experiment, each column to a pathway gene. The entry in row e and column g is equal 1 if experiment e targets gene g.
Gs	Optional list of graphs (as matrices) to search the best from. Default = NULL.
inference	search to use exhaustive enumeration over provided set of graphs or over all possible, greedy for greedy hillclimbing
parameter.estimation	linear.reg (default) for linear regression, bayes.linear.reg for bayesian linear regression.
verbose	Allows printout of progression statements. Default: FALSE

Details

For the method to provably work, we need at least one perturbation experiment per each single gene, and at least one experiment per each pair. Check your experiments matrix E! The effect measurements can be any real vectors. Bayesian linear regression (for parameter.estimation = "bayes.linear.reg") takes much longer to compute than linear regression (parameter.estimation = "linear.reg").) The function plot.lem plots the inferred pathway as a directed graph, and the scores of all models evaluated in the search for the model that best fits the data.

Value

graph	inferred directed pathway gene graph (graphNEL object)
score	(marginal) log likelihood of model(s). In the case when bayesian regression was used for inference, it is the marginal log likelihood. In the case when linear regression was used, it is log likelihood.
beta	estimate of the vector of contributions of the genes in the pathway graph to the phenotype
sigma	standard deviation of the phenotype
b	For parameter.estimation = "bayes.linear.reg" it is the prior parameter b. By default, for parameter.estimation = "linear.reg", b is NULL.
all.scores	vector of scores of all models considered in the inference of the pathway graph.
inference	Method used for infering the graph. Either "greedy" for greedy search or "search" for exhaustive search over user-defined set of graphs or over all possible graphs of given size.
parameter.estimation	Method used for estimating model parameters and scoring the model. Either "linear.reg" for linear regression or "bayes.linear.reg" for bayesian linear regression.

Author(s)

Ewa Szczurek

References

Szczurek, E & Beerenwinkel, N., Linear Effects Models from Combinatorial Perturbation Data. Bioinformatics, 2016

See Also

plot.lem

Examples

```
set.seed(111)

  sim.set <- simulateLEM(n=5, b = 1, sigma = 0.01, rep.no = 5, Y.no=1)
 lem.res <- lem( sim.set$Y, sim.set$E, inference = "greedy", parameter.estimation = "bayes.linear.reg",verbo
 print(sim.set$G)
 print.lem(lem.res)
 plot.lem(lem.res)
```

---

plot.lem *Linear Effects Models - plotting*

Description

Plots the graph or the scores of all models which were searched through

Usage

```
plot.lem(x, what="graph", remove.singletons=FALSE, PDF=FALSE, filename="lemplot.pdf",  transitive
```

Arguments

x	Lem object (result of calling lem)
what	What to plot. (i) "graph" - the pathway graph (default), (ii) "scores" - the sorted scores of all models that were evaluated to identify the model which best fits the data.
remove.singletons	Remove unconnected, singleton nodes from the plot of the pathway graph. Valid for what="graph".
PDF	Save a pdf file with the plot.
filename	A name of the pdf file.
transitiveReduction	Plot only a transitive reduction of the identified graph. Valid for what="graph".
...	other arguments to be passed to the Rgraphviz plot function or to the graphics 'plot' function.

Value

none

Author(s)

Ewa Szczurek

See Also

lem

`print.lem` *Linear Effects Models - printing basic info*

Description

Prints the Linear Effects Model parameters

Usage

```
print.lem(x)
```

Arguments

x	Lem object (result of calling lem)

Details

This function will print the adjacency matrix of the graph, the beta coefficients, the score, and the inference and parameter estimation methods used for learning.

Value

none

Author(s)

Ewa Szczurek

See Also

lem

simulateLEM *Simulate a Linear Effects Model*

Description

Generates random graph, experiments matrix, beta contributions and data (measurements)

Usage

```
simulateLEM(n, b = 1, sigma = 0.01, rep.no = 3, Y.no=1, collapse=FALSE)
```

Arguments

n	Number of genes perturbed in experiments (the genes in the pathway modeled by LEM).
b	The value for the hyper-prior parameter b of the distribution of the beta parameters. We assume the betas distribution is a zero mean isotropic Gaussian with precision b, By default, b=1.
sigma	The standard deviation for the Gaussian noise around the simulated measurements. By default, sigma=0.01.
rep.no	The number of times each experiment is repeated.
Y.no	The number of measurement vectors to simulate.

Details

For the LEM to be identifiable, the experiments have to at least perturb each single and each pair of nodes. The generated experiment matrix E has rep.no * n * (n choose 2) rows. The rows correspond to experiments and columns to perturbed genes.

For Y.no = 1, one vector of beta coefficients will be sampled, and one vector of measurements (one vector entry for each experiment) will be generated according to the model. For Y.no >1, a matrix of coefficients will be sampled, and one matrix of measurements will be returned, with one column per one measurement vector. This corresponds to assuming that the pathway is always the same, but if a number of phenotypes is measured, each of them will be different and will have different gene contributions.

The function returns a list.

Value

Y	The simulated measurement vector (matrix for Y.no>1).
betas	The simulated gene contributions vector (matrix for Y.no>1)
E	The experiment matrix.
G	The simulated graph as an adjacency matrix.

Author(s)

Ewa Szczurek

Index