Package 'lapmix'

March 31, 2025

Version 1.72.0

Title Laplace Mixture Model in Microarray Experiments

Author Yann Ruffieux, contributions from Debjani Bhowmick, Anthony C. Davison, and Darlene R. Goldstein

Maintainer Yann Ruffieux <yann.ruffieux@epfl.ch>

Description Laplace mixture modelling of microarray experiments. A hierarchical Bayesian approach is used, and the hyperparameters are estimated using empirical Bayes. The main purpose is to identify differentially expressed genes.

Depends R (>= 2.6.0),stats

Imports Biobase, graphics, grDevices, methods, stats, tools, utils

License GPL (>= 2)

URL http://www.r-project.org, http://www.bioconductor.org, http://stat.epfl.ch

biocViews Microarray, OneChannel, DifferentialExpression

git_url https://git.bioconductor.org/packages/lapmix

git_branch RELEASE_3_20

git_last_commit 47b3250

git_last_commit_date 2024-10-29

Repository Bioconductor 3.20

Date/Publication 2025-03-31

Contents

---

lap.volcanoplot   *Laplace Volcano Plot*

Description

Creates a volcano plot of log-fold changes versus log-odds of differential expression under the Laplace mixture model.

Usage

```
lap.volcanoplot(res, highlight=0, ...)
```

Arguments

res   the output from the `lapmix.Fit` routine

highlight   number of genes to be highlighted; genes are highlighted in descending order of their posterior odds

...   additional arguments given to the plot function

Details

A volcano plot is any plot which displays fold changes versus a measure of statistical significance of the change.

Value

A plot is created on the current graphics device.

Author(s)

Yann Ruffieux

Examples

```
#  See lapmix.Fit example
```

---

lapmix-internal   *Internal Lapmix Functions*

Description

Internal Lapmix Functions

Details

These routines are for internal use only!

---

lapmix.Fit	*Empirical Bayes Statistics for Differential Expression under Laplace Model*

Description

Computes posterior odds of differential expression under the Laplace mixture model, with parameters estimated using an empirical Bayes approach.

Usage

```
lapmix.Fit(Y, asym=FALSE, fast=TRUE, two.step=TRUE,
           w=0.1, V=10, beta=0, gamma=1, alpha=0.1)
```

Arguments

Y	data.frame or matrix containing the log relative expression levels, where each row represents a gene. Alternatively, a list of arrays of possibly different sizes, or an object of class `eSet` or `ExpressionSet`.
asym	indicates whether the asymmetric Laplace model is used rather than the symmetric one.
fast	indicates whether the 'fast' estimation method is used.
two.step	indicates whether the two-step estimation method is used; otherwise the marginal likelihood is maximised in one step.
w, V, beta, gamma, alpha	
	initial values given to the optimization algorithms for the estimation of the hyperparameters.

Details

This method fits the results of a microarray experiment to a Laplace mixture model. These results are assumed to take the form of normalized base 2 logarithm of the expression ratios. An empirical Bayes approach is used to estimate the hyperparameters of the model. The `lap.lodds` is sometimes known as the L-statistic (if the symmetric model is used) or the AL-statistic (if the asymmetric model is used). These statistics can be used to rank the genes according to the posterior odds of differential expression, via the routine `laptopTable`. They can be visualized using the `lap.volcanoplot` function.

If there are different numbers of replicates between genes, one may wish to write the data in a list of arrays. If a matrix representation is desired, one can stick in NaN's where appropriate.

The 'fast' estimation method ignores the integrals which cannot be computed with the t-distribution function. This method is suggested, since these problematic integrals are few and far between. The estimates are practically not affected, and we avoid the potential problems that arise when integrating numerically with the `integrate` function.

Value

lap.lods	numeric vector containing the posterior log-odds of differential expression
prob	numeric vector containing the posterior probabilities of differential expression
med.number	number of differentially expressed genes according to the median rule

M	numeric vector with average log fold changes within genes
s_sq	numeric vector with sample variances within genes
nb.rep	numeric vector with number of replicates within genes
estimates	list containing the empirical Bayes estimates of the hyperparameters
code	integer indicating why the likelihood optmization terminated, cf. nlm routine. There are two such indicators in the case of the two-step estimation.

Author(s)

Yann Ruffieux

References

Bhowmick, D., Davison, A.C., and Goldstein, D.R. (2006). A Laplace mixture model for identification of differential expression in microarray experiments.

Examples

```
# Simulate gene expression data under Laplace mixture model: 3000 genes with
# 4 duplicates each; one gene in ten is differentially expressed.

G <- 3000
Y <- NULL
sigma_sq <- 1/rgamma(G, shape=2.8, scale=0.04)
mu <- rexp(G, rate=1/(sigma_sq*1.2))-rexp(G, rate=1/(sigma_sq*1.2))
is.diff <- sample(c(0,1), replace=TRUE, prob=c(0.9,0.1), size=G)
mu <- mu*is.diff
for(g in 1:G)
 Y <- rbind(Y, rnorm(4,mu[g], sd=sqrt(sigma_sq[g])))

# with symmetric model
res <- lapmix.Fit(Y)
res$estimates
laptopTable(res, 20)
lap.volcanoplot(res, highlight=res$med.number)

# with asymmetric model
res2 <- lapmix.Fit(Y, asym=TRUE)
res2$estimates
laptopTable(res2, 20)
lap.volcanoplot(res2, highlight=res2$med.number)
```

---

laptopTable	*Table of Top Genes from Laplace Linear Model Fit*

Description

Extract a table of the top-ranked genes from a Laplace mixture model fit.

Usage

```
laptopTable(res, number=res$med.number, sort.by='L')
```

Arguments

res	the output from the `lapmix.Fit` routine
number	how many genes to pick out; if missing: number is determined by the median rule
sort.by	character string specifying statistic to sort the selected genes by in the output data.frame.

Details

This function summarizes a Laplace mixture model fit object produced by `lapmix.Fit` by selecting the top-ranked genes according to the posterior log-odds or M-values.

The `sort.by` argument specifies the criterion used to select the top genes. Only two choices at the moment: "M" to sort by the (absolute) coefficient representing the log-fold-change, and "L" to sort by the posterior odds of differential expression under the Laplace mixture model.

Value

A dataframe with a row for the `number` top genes and the following columns:

M	average log fold change
log.odds	log posterior odds that the gene is differentially expressed

Author(s)

Yann Ruffieux

Examples

```
#  See lapmix.Fit example
```

Index