Package ‘ORCME’

July 31, 2015

Type Package
Title Order Restricted Clustering for Microarray Experiments
Version 2.0.2
Date 2015-07-15
Author Adetayo Kasim, Martin Otava, Tobias Verbeke
Maintainer Rudradev Sengupta <rudradev.sengupta@uhasselt.be>
Description Provides clustering of genes with similar
dose response (or time course) profiles. It implements the method
described by Lin et al. (2012).
Imports Iso
License GPL-3
LazyLoad yes
Repository CRAN
Repository/R-Forge/Project orcme
Repository/R-Forge/Revision 65
Repository/R-Forge/DateTimeStamp 2015-07-23 12:31:52
Date/Publication 2015-07-31 12:12:01
NeedsCompilation no
Depends R (>= 2.10)

R topics documented:

doseData ................................................................. 2
geneData .............................................................. 2
monotoneDirection .................................................... 3
ORCME ................................................................. 5
plotCluster ............................................................. 6
plotIsomeans .......................................................... 7
plotLambda ............................................................. 9
resampleORCME ...................................................... 10

Index 12
### doseData

**Dose Data Example**

**Description**
Dose data; a vector of length 12 with 3 observations for each of 4 doses.

**Usage**
```
data(doseData)
```

**Format**
The format is: num [1:12] 1 1 1 2 2 2 3 3 3 ...

**Examples**
```
data(doseData)
doseData
```

### geneData

**Gene Expression Data Example**

**Description**
This dose-response microarray data contains 1000 genes and 4 doses (one control dose (zero dose) and three increasing dose) with 3 arrays at each dose level. Due to confidentiality, it is only part of the real data set.

**Usage**
```
data(geneData)
```

**Format**
A data frame with 1000 observations on the following 12 variables.

- X1  Sample one with zero dose
- X1.1 Sample two with zero dose
- X1.2 Sample three with zero dose
- X2  Sample one with second dose
- X2.1 Sample two with second dose
monotoneDirection

X2.2 Sample three with second dose
X3 Sample one with third dose
X3.1 Sample two with third dose
X3.2 Sample three with third dose
X4 Sample one with fourth dose
X4.1 Sample two with fourth dose
X4.2 Sample three with fourth dose

References


Examples

data(geneData)
head(geneData)

monotoneDirection gene expression matrix for all genes

The monotone means under increasing/decreasing trend

Description

The function calculates the likelihood for the increasing and decreasing trend in the dose response for all the given genes separately gene-by-gene. The trend with the higher likelihood is chosen and the isotonic regression is applied on the means.

Usage

monotoneDirection(geneData, doseData)

Arguments

geneData gene expression matrix for all genes
doseData indicates the dose levels
monotoneDirection

Value

A list with components

direction the direction with the higher likelihood of increasing (indicated by "up") or decreasing (indicated by "dn") trend.

incData isotonic means with respect to dose for those genes that were classified as following the increasing trend.

decData isotonic means with respect to dose for those genes that were classified as following the decreasing trend.

obsincData observed gene expression matrix for those genes that were classified as following the increasing trend.

obsdecData observed gene expression matrix for those genes that were classified as following the decreasing trend.

arrayMean isotonic means with respect to dose for all genes.

Author(s)

Adetayo Kasim, Martin Otava and Tobias Verbeke

References


See Also

ORCME, plotIsomeans

Examples

data(doseData)
data(geneData)

dirData <- monotoneDirection(geneData = geneData, doseData = doseData)

## direction of monotone trend
Direction <- dirData$direction

## isotonic means for upward genes
incData <- as.data.frame(dirData$incData)

## isotonic means for downward genes
decData <- as.data.frame(dirData$decData)

## observed data upward genes
obsIncData <- as.data.frame(dirData$obsIncData)

## observed data for downward genes
obsDecData <- as.data.frame(dirData$obsDecData)

## isotonic means for all genes
isoMeans <- as.data.frame(dirData$arrayMean)
Description

The function performs delta-clustering of a microarray data. It can be used for clustering of both the time-course or dose-response microarray data.

Usage

\texttt{ORCME}(<\texttt{DRdata}, \texttt{lambda}, \texttt{phi}, \texttt{robust=FALSE})

Arguments

\begin{itemize}
  \item \texttt{DRdata} matrix of a microarray data with rows corresponding to genes and columns corresponding to time points or different doses
  \item \texttt{lambda} assumed proportion of coherence relative to the observed data, it ranges between 0 and 1. A lambda value of 1 considers the observed data as a cluster and lambda value of 0 finds every possible pattern within the data.
  \item \texttt{phi} minimum number of genes in a cluster
  \item \texttt{robust} logical variable that determines, if algorithm uses robust version based on median polish and absolute values, instead of mean square error. Default is \texttt{FALSE}.
\end{itemize}

Value

The matrix of classification into clusters: each row represents one gene and columns found clusters. The matrix consist of the Booleans values, in each row there is only one of them \texttt{TRUE} which means that the gene was classified into the respective cluster.

Author(s)

Adetayo Kasim, Martin Otava and Tobias Verbeke

References


See Also

\texttt{monotoneDirection, plotIsomeans}
Examples

```r
data(doseData)
data(geneData)

dirData <- monotoneDirection(geneData = geneData, doseData = doseData)
incData <- as.data.frame(dirData)$incData

print(orcmefull <- ORCME(DRdata=incData, lambda=0.15, phi=2))
orcmefullRobust <- ORCME(DRdata=incData, lambda=0.15, phi=2, robust=TRUE)

# number of genes within cluster
colSums(orcmefull)
colSums(orcmefullRobust)
```

plotCluster

**Plotting the gene specific profiles for one given cluster of genes**

Description

The function is plotting the profiles of the genes that belong to the same cluster. It is not providing the clustering itself, just plotting the results of clustering from input. Optionally, the function can center the profiles around the gene-specific means.

Usage

```r
plotCluster(DRdata, doseData, ORCMEoutput, clusterID, zeroMean=FALSE, xlabel, ylabel, main="")
```

Arguments

- **DRdata**
  - the microarray data with rows corresponding to genes and columns corresponding to time points or different doses
- **doseData**
  - indicates the dose levels
- **ORCMEoutput**
  - the matrix of classification into clusters: each row represents one gene and columns found clusters. The matrix consist of the Booleans values, in each row there is only one of them TRUE which means that the gene was classified into the respective gene
- **clusterID**
  - id of the cluster to be plotted
- **zeroMean**
  - if TRUE, it centers the gene profiles around the gene-specific means, default is FALSE
- **xlabel**
  - a title for the x axis
- **ylabel**
  - a title for the y axis
- **main**
  - an overall title for the plot
Value

Plot of the gene specific profiles dependent one the dose level (or time point) that are classified into the given cluster.

Author(s)

Adetayo Kasim, Martin Otava and Tobias Verbeke

References


See Also

ORCME, plotIsomeans

Examples

data(doseData)
data(geneData)

dirData <- monotoneDirection(geneData = geneData,doseData = doseData)
incData <- as.data.frame(dirData$incData)
ORCMEoutput <- ORCME(DRdata=incData,lambda=0.1,phi=2)

plotCluster(DRdata=incData,doseData=doseData, ORCMEoutput=ORCMEoutput, clusterID=4,zeroMean=FALSE, xlab="Dose",ylab="Gene Expression")

plotIsomeans

Plot of the observed gene expression and the isotonic means with respect to dose

Description

The function is plotting the observed data points of the gene expression and isotonic means with respect to dose for one particular gene.

Usage

plotIsomeans(monoData, obsData, doseData, geneIndex)
Arguments

- `monoData`: isotonic means with respect to dose for all genes
- `obsData`: observed gene expression for all genes
- `doseData`: indicates the dose levels
- `geneIndex`: index of the gene to be plotted

Value

Plot of the data points and the isotonic means for each dose with the isotonic regression curve.

Author(s)

Adetayo Kasim, Martin Otava and Tobias Verbeke

References


See Also

`orcme`, `monotoneDirection`

Examples

data(dosedata)
data(genedata)
dirData <- monotoneDirection(geneData = genedata, doseData = dosedata)
inData <- as.data.frame(dirData$incData)
obsIncData <- as.data.frame(dirData$obsincData)

## gene-specific profile plot
plotIsomeans(monoData=incData, obsData=obsIncData, doseData=
doseData, geneIndex=10)
plotLambda

Plot the variety of the properties dependent on the proportion of heterogeneity in observed data set

Description

This function provides the plots of the dependency of the variety of properties on the proportion of heterogeneity in observed data set. It is not using the clustering as simple input, but it is also computing additional properties. The function can plot within cluster sum of squares, number of cluster, penalized within cluster sum of squares, Calsanzik and Harabasx index and Hartigan index.

Usage

plotLambda(lambdaChoiceOutput, output)

Arguments

lambdaChoiceOutput

the output of the function resampleORCME

output

the variable that determines which output would be plotted, the values are "wss" for the cluster sum of squares, "ncluster" for the number of cluster, "pwss" for the penalized within cluster sum of squares, "ch" for the Calsanzik and Harabasx index and "h" for the Hartigan index

Value

A plot of one of the properties mentioned above dependent on the proportion of heterogeneity. The confidence intervals are plotted instead of the point estimates.

Author(s)

Adetayo Kasim, Martin Otava and Tobias Verbeke

References


See Also

ORCME, resampleORCME
Examples

data(doseData)
data(geneData)
dirData <- monotoneDirection(geneData = geneData, doseData = doseData)
incData <- as.data.frame(dirData$incData)

lambdaVector <- c(0.05, 0.50, 0.95)

lambdaChoiceOutput <- resampleORCME(clusteringData=incData, lambdaVector=lambdaVector)
plotLambda(lambdaChoiceOutput, output="wss")
plotLambda(lambdaChoiceOutput, output="ncluster")
plotLambda(lambdaChoiceOutput, output="pwss")
plotLambda(lambdaChoiceOutput, output="ch")
plotLambda(lambdaChoiceOutput, output="h")

resampleORCME

Estimation of the proportion of the heterogeneity in the observed data for clustering

Description

The function is computing within cluster sum of squares for given proportion of heterogeneity. Minimal number of genes per cluster is fixed as 2. The sum of squares is computed through resampling the 100 data sets with 100 genes randomly sampled with replacement from the reduced expression data.

Usage

resampleORCME(clusteringData, lambdaVector, robust=FALSE)

Arguments

clusteringData the microarray data with rows corresponding to genes and columns corresponding to time points or different doses

lambdaVector vector of assumed proportions of of heterogeneity of the observed data, it ranges between 0 and 1. A lambda value of 1 considers the observed data as a cluster and lambda value of 0 finds every possible pattern within the data

robust logical variable that determines, if algorithm uses robust version based on median polish and absolute values, instead of mean square error. Default is FALSE.
Value
A list of matrices that represent one of the 100 iterations. Every matrix consist of the columns
lambda vector of the proportions of heterogeneity given as input
WSS within clusters sum of squares for given proportion of heterogeneity
TSS total clusters sum of squares for given proportions of heterogeneity
nc number of clusters as a function for given proportions of heterogeneity

Author(s)
Adetayo Kasim, Martin Otava and Tobias Verbeke

References
International Conference on Intelligent Systems for Molecular Biology, 1, 93-103.

See Also
orcme, plotLambda

Examples
```r
data(doseData)
data(geneData)
dirData <- monotoneDirection(geneData = geneData, doseData = doseData)
incData <- as.data.frame(dirData$incData)
lambdaVector <- c(0.05, 0.50, 0.95)

resampleORCME(clusteringData=incData, lambdaVector=lambdaVector, robust=FALSE)
```
Index

*Topic **cluster**
  monotoneDirection, 3
  ORCME, 5
  plotCluster, 6
  plotIsomeans, 7
  plotLambda, 9
  resampleORCME, 10

*Topic **datasets**
  doseData, 2
  geneData, 2

doseData, 2

geneData, 2

monotoneDirection, 3, 5, 8

ORCME, 4, 5, 7–9, 11

plotCluster, 6
plotIsomeans, 4, 5, 7, 7
plotLambda, 9, 11

resampleORCME, 9, 10