Package 'nnNorm'

October 18, 2022

Version 2.60.0

Date 2010-04-13

Title Spatial and intensity based normalization of cDNA microarray data based on robust neural nets

Author Adi Laurentiu Tarca <atarca@med.wayne.edu>

Depends R(>= 2.2.0), marray

Imports graphics, grDevices, marray, methods, nnet, stats

Maintainer Adi Laurentiu Tarca <atarca@med.wayne.edu>

Description This package allows to detect and correct for spatial and intensity biases with two-channel microarray data.
The normalization method implemented in this package is based on robust neural networks fitting.

biocViews Microarray, TwoChannel, Preprocessing

License LGPL

URL http://bioinformaticsprb.med.wayne.edu/tarca/

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

git_branch RELEASE_3_15

git_last_commit 7b10a62

git_last_commit_date 2022-04-26

Date/Publication 2022-10-18

R topics documented:

compNorm — *Compares the distribution of several vectors at a time using either boxplots or density curves*

Description

This function was concieved to easily compare several normalization methods in terms of variability of log-ratios, M. Basically it produces two plots: The first is a the density plot of the several matrices passed as arguments, while the second is a box plot. Median of absolute deviations for each method is printed on screen.

Usage

```
compNorm(x,...,bw="AUTO",xlim=c(-3,3),titles="AUTO",type="d")
```

Arguments

x	A vector of numerical values, e.q. the M values of a data set: `as.vector(maM(swirl))`.
...	An undefined number of objects similar with x.
bw	Band width required to compute the density distribution. "AUTO" will adjust bw to a suitable value.
xlim	The range for abscissa of the density plots.
titles	Names to be displayed the charts legend. "AUTO" will use the matrices names passed as arguments. .
type	If set to "d", density plot will be shown; if set to "d" box plot will be shown.

Details

This function is used to compare the normalized log ratios M obtained with several normalization methods.

Value

NULL, this function only displays charts and prints on the screen some statistics.

Author(s)

Tarca, A.L.

References

A. L. Tarca, J. E. K. Cooke, and J. Mackay. Robust neural networks approach for spatial and intensity dependent normalization of cDNA data. Bioinformatics. 2004,submitted.

See Also

maNormNN

Examples

```
# Normalize swirl data with two methods
data(swirl)
swirlNN<-maNormNN(swirl[,1])
swirlLoess<-maNormMain(swirl[,1])
nms<-c("None","Loess","NNets")
#compare distributions: density plot
compNorm(as.vector(maM(swirl[,1])),as.vector(maM(swirlLoess)),as.vector(maM(swirlNN)),xlim=c(- 2,2),bw="AUTO",
#compare distributions: box plot
compNorm(as.vector(maM(swirl[,1])),as.vector(maM(swirlLoess)),as.vector(maM(swirlNN)),xlim=c(- 2,2),bw="AUTO",
```

---

detectSpatialBias *Detecting spatial bias within the print-tips of a two channel array*

Description

This function allows to identify in two channel batch of arrays, which are the print-tips where spatial bias is present.

Usage

```
detectSpatialBias(mbatch, corThreshold=0.6)
```

Arguments

mbatch	A marrayRaw or marrayNorm batch of two channel arrays.
corThreshold	The correlation treshold to be used.

Details

This function computes two matrices: biasRow and biasCol. The elements of these matrices represent the fraction of rows (columns) for which the correlation coefficient between log-ratios, M, and column index (row index) is higher than a user specified treshold (default corThreshold=0.6). The idea here is to see in which print-tip a important fraction of the rows (columns) are highly correlated with the column (row) index. Since some rows (columns) will show positive correlation while the other negative correlation, we are only interested in a sigle direction of the correlation, i.e. either positive or negative.

Value

This function returns a list with two matrices. biasRow and biasCol. The rows of these matrices correspond to the print tips counted metaRow wise, and the columns correspond to arrays. Values in these matrices superior to 33 point to print-tips that have more tha a third of the rows (columns) with important spatial bias.

Author(s)

Tarca, A.L.

References

A robust neural networks approach for spatial and intensity dependent normalization of cDNA microarray data, Adi. L. Tarca , Janice. E. K. Cooke, and John Mackay, Bioinformatics, 21, 2005, 2674 - 2683.

See Also

maNormNN

Examples

```
# detecting spatial bias in swirl data
data(swirl)
# print-tip, intensity and spatial normalization of the first slide in swirl data set
myres<-detectSpatialBias(swirl)
```

---

maNormNN *Intensity and spatial normalization using robust neural networks fitting*

Description

This function normalizes a batch of cDNA arrays by removing the intensity and spatial dependent bias.

Usage

```
maNormNN(mbatch,w=NULL,binWidth=3,binHeight=3,model.nonlins=3,iterations=100,nFolds=10,maplots=FALS
```

Arguments

mbatch	A marrayRaw or marrayNorm batch of arrays.
w	Weights to be assigned to each spot. If provided, it should be a vector with the same length as maNspots(mbatch).
binWidth	Width of the bins in the X direction (spot column) in which the print tip will be divided in order to account for spatial variation. Max value is maNsc(mbatch), Min value is 1. However if it is set to a number larger than maNsc(mbatch)/2 (so less than two bins in X direction) the variable X will not be used as predictor to estimate the bias.

binHeight	Height of the bins in the Y direction (spot row)in which the print tip will be divided in order to account for spatial variation. Max value is maNsr(mbatch), Min value is 1. However if it is set to a number larger than maNsr(mbatch)/2 (so less than two bins in Y direction) the variable Y will not be used as predictor to estimate the bias.
model.nonlins	Number of nodes in the hidden layer of the neural network model.
iterations	The number of iterations at which (if not converged) the training of the neural net will be stopped.
nFolds	Number of cross-validation folds. It represents the number of equal parts in which the data from a print tip is divided into: the model is trained on nFolds-1 parts and the bias is estimated for one part at the time. Higher values improve the results but increase the computation time. Ideal values are between 5 and 10.
maplots	If set to "TRUE" will produce a $M - A$ plot for each slide before and after normalization.
verbose	If set to "TRUE" will show the output of the nnet function which is training the neural network models.

Details

This function uses neural networks to model the bias in cDNA data sets.

Value

A marrayNorm object containing the normalized log ratios. See marrayNorm class for details

Author(s)

Tarca, A.L.

References

A. L. Tarca, J. E. K. Cooke, and J. Mackay. Robust neural networks approach for spatial and intensity dependent normalization of cDNA data. Bioinformatics. 2004,submitted.

See Also

compNorm,nnet

Examples

```
# Normalization of swirl data
data(swirl)
# print-tip, intensity and spatial normalization of the first slide in swirl data set
swirlNN<-maNormNN(swirl[,1])

#do not consider spatial variations, and display M-A plots before and after normalization
swirlNN<-maNormNN(swirl[,1],binWidth=maNsc(swirl),binHeight=maNsr(swirl),maplots=TRUE)
```

Index