Package 'qsmooth'

October 17, 2020

Title Smooth quantile normalization

Version 1.4.0

Imports SummarizedExperiment, utils, sva, stats, methods, graphics

Depends R (>= 4.0)

Suggests bodymapRat, quantro, knitr, rmarkdown, BiocStyle, testthat

VignetteBuilder knitr

RoxygenNote 6.1.1

Encoding UTF-8

License CC BY 4.0

Description Smooth quantile normalization
is a generalization of quantile normalization,
which is average of the two types of assumptions
about the data generation process: quantile
normalization and quantile normalization between
groups.

biocViews Normalization, Preprocessing, MultipleComparison,
Microarray, Sequencing, RNASeq, BatchEffect

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

git_branch RELEASE_3_11

git_last_commit f5ee169

git_last_commit_date 2020-04-27

Date/Publication 2020-10-16

Author Stephanie C. Hicks [aut, cre] (<https://orcid.org/0000-0002-7858-0231>),
Kwame Okrah [aut],
Hector Corrada Bravo [aut] (<https://orcid.org/0000-0002-1255-4444>),
Rafael Irizarry [aut] (<https://orcid.org/0000-0002-3944-4309>)

Maintainer Stephanie C. Hicks <shicks19@jhu.edu>

R topics documented:

1

qsmooth	*qsmooth*

Description

This function applies a generalization of quantile normalization called smoothed quantile normalization. This function defines the qsmooth class and constructor.

Usage

```
qsmooth(object, group_factor, batch = NULL, norm_factors = NULL, window = 0.05)
```

Arguments

object   an object which is a `matrix` or `data.frame` with observations (e.g. probes or genes) on the rows and samples as the columns. Alternatively, a user can provide a `SummarizedExperiment` object and the `assay(object,"counts")` will be used as input for the qsmooth normalization.

group_factor   a group level continuous or categorial covariate associated with each sample or column in the `object`. The order of the `group_factor` must match the order of the columns in `object`.

batch   (Optional) batch covariate (multiple batches are not allowed). If batch covariate is provided, `Combat()` from `sva` is used prior to qsmooth normalization to remove batch effects. See `Combat()` for more details.

norm_factors   optional normalization scaling factors.

window   window size for running median which is a fraction of the number of rows in `object`. Default is 0.05.

Details

Quantile normalization is one of the most widely used normalization tools for data analysis in genomics. Although it was originally developed for gene expression microarrays it is now used across many different high-throughput applications including RNAseq and ChIPseq. The methodology relies on the assumption that observed changes in the empirical distribution of samples are due to unwanted variability. Because the data is transformed to remove these differences it has the potential to remove interesting biologically driven global variation. Therefore, applying quantile normalization, or other global normalization methods that rely on similar assumptions, may not be an appropriate depending on the type and source of variation.

This function computes a weight at every quantile that compares the variability between groups relative to within groups. In one extreme quantile normalization is applied and in the other extreme quantile normalization within each biological condition is applied. The weight shrinks the group-level quantile normalized data towards the overall reference quantiles if variability between groups is sufficiently smaller than the variability within groups. See the vignette for more details.

Value

A object of the class qsmooth that contains a numeric vector of the qsmooth weights in the `qsmoothWeights` slot and a matrix of normalized values after applying smoothed quantile normalization in the `qsmoothData` slot.

Examples

```
dat <- cbind(matrix(rnorm(1000), nrow=100, ncol=10),
             matrix(rnorm(1000, .1, .7), nrow=100, ncol=10))
dat_qs <- qsmooth(object = dat,
                  group_factor = rep(c(0,1), each=10))
```

---

qsmooth-class   *the qsmooth class*

Description

Objects of this class store all the values needed information to work with a qsmooth object

Value

qsmoothWeights returns the qsmooth weights and qsmoothData returns the qsmooth normalized data

Slots

qsmoothWeights qsmooth weights

qsmoothData qsmooth normalized data

Examples

```
dat <- cbind(matrix(rnorm(1000), nrow=100, ncol=10),
             matrix(rnorm(1000, .1, .7), nrow=100, ncol=10))
dat_qs <- qsmooth(object = dat,
                  group_factor = rep(c(0,1), each=10))
```

---

qsmoothData   *Generic function that returns the qsmooth normalized data*

Description

Given a qsmooth object, this function returns the qsmooth normalized data

Accessors for the 'qsmoothData' slot of a qsmooth object.

Usage

```
qsmoothData(object)

## S4 method for signature 'qsmooth'
qsmoothData(object)
```

Arguments

object   an object of class qsmooth.

Value

The normalized data after applying smoothed quantile normalization.

Examples

```
dat <- cbind(matrix(rnorm(1000), nrow=100, ncol=10),
             matrix(rnorm(1000, .1, .7), nrow=100, ncol=10))
dat_qs <- qsmooth(object = dat,
                  group_factor = rep(c(0,1), each=10))
qsmoothData(dat_qs)
```

---

qsmoothPlotWeights   *Plot weights from* qsmooth *function.*

Description

This function plots a scatterplot showing the qsmoothWeights along the y-axis and the quantiles on the x-axis.

Usage

```
qsmoothPlotWeights(
  object,
  xLab = "quantiles",
  yLab = "weights",
  mainLab = "qsmooth weights"
)
```

Arguments

object	a qsmooth object from qsmooth
xLab	label for x-axis. Default is "quantiles"
yLab	label for y-axis. Default is "weights"
mainLab	title of plot. Default is "qsmooth weights"

Value

A scatterplot will be created showing the qsmoothWeights along the y-axis and the quantiles on the x-axis.

Examples

```
dat <- cbind(matrix(rnorm(1000), nrow=100, ncol=10),
             matrix(rnorm(1000, .1, .7), nrow=100, ncol=10))
dat_qs <- qsmooth(object = dat,
                  group_factor = rep(c(0,1), each=10))
qsmoothPlotWeights(dat_qs)
```

---

qsmoothWeights *Generic function that returns the qsmooth weights*

Description

Given a qsmooth object, this function returns the qsmooth weights

Accessors for the 'qsmoothWeights' slot of a qsmooth object.

Usage

```
qsmoothWeights(object)

## S4 method for signature 'qsmooth'
qsmoothWeights(object)
```

Arguments

object   an object of class qsmooth.

Value

The weights calculated for each feature after applying smoothed quantile normalization.

Examples

```
dat <- cbind(matrix(rnorm(1000), nrow=100, ncol=10),
             matrix(rnorm(1000, .1, .7), nrow=100, ncol=10))
dat_qs <- qsmooth(object = dat,
                  group_factor = rep(c(0,1), each=10))
qsmoothWeights(dat_qs)
```

---

qstats *qstats*

Description

This function is a helper function that computes quantile statistics for the function qsmooth.

Usage

```
qstats(object, group_factor, window = 0.05)
```

Arguments

object   an object which is a data frame or matrix with observations (e.g. probes or genes) on the rows and samples as the columns.

group_factor   a group level continuous or categorial covariate associated with each sample or column in the object. The order of the group_factor must match the order of the columns in object.

window   window size for running median which is a fraction of the number of rows in object. Default is 0.05.

Value

A list of quantile statistics including

Q	sample quantiles
Qref	reference quantile
Qhat	linear model fit at each quantile
SST	total sum of squares
SSB	between sum of squares
SSE	within sum of squares
roughWeights	SSE / SST
smoothWeights	smoothed weights computed using a running median with a given window size.

Examples

```
dat <- cbind(matrix(rnorm(1000), nrow=100, ncol=10),
             matrix(rnorm(1000, .1, .7), nrow=100, ncol=10))
qs <- qstats(object = dat,
             group_factor = rep(c(0,1), each=10),
             window = 0.05)
```

Index