Package 'RcmdrPlugin.depthTools'

January 20, 2025

Type Package

Title R Commander Depth Tools Plug-in

Version 1.4

Date 2020-10-15

Author Sara Lopez-Pintado <sl2929@columbia.edu> and Aurora Torrente
<etorrent@est-econ.uc3m.es>.

Maintainer Aurora Torrente <etorrent@est-econ.uc3m.es>

Depends Rcmdr (>= 1.4-0), tcltk, depthTools

Description We provide an Rcmdr plug-in based on the depthTools
package, which implements different robust statistical tools
for the description and analysis of gene expression data based
on the Modified Band Depth, namely, the scale curves for
visualizing the dispersion of one or various groups of samples
(e.g. types of tumors), a rank test to decide whether two
groups of samples come from a single distribution and two
methods of supervised classification techniques, the DS and
TAD methods.

License GPL (>= 2)

NeedsCompilation no

Repository CRAN

Date/Publication 2020-10-16 04:30:02 UTC

Contents

RcmdrPlugin.depthTools-package
Graphical User Interface for depthTools

Description

This package provides an Rcmdr "plug-in" based on the depthTools package. It provides a GUI for the computation and graphical visualization of the Modified Band Depth (MBD) of points in a data set with respect to the same or a different set. It also allows to compute the scale curve of a set, to perform the rank test to decide whether two samples come from the same population, and to classify new data points according to the DS and the TAD classification methods. In addition, the user can obtain a plot which enhances a fixed percentage of the most central curves through the centralPlot, and compute a trimmed mean, based on MBD.

Details

Package: RcmdrPlugin.dephTools Type: Package Version: 1.2 Date: 2013-02-14 License: GPL (>= 2)

Author(s)

Aurora Torrente <etorrent@est-econ.uc3m.es> and Sara Lopez-Pintado <sl2929@columbia.edu>

Maintainer: Aurora Torrente <etorrent@est-econ.uc3m.es>

See Also

[Rcmdr.](Rcmdr)

---

computeMBD   *Computation of the Modified Band Depth in R-commander*

Description

computeMBD computes the Modified Band Depth of a sample, the active data set, either with respect to the same data set or a different one.

Details

The MBD of the active data set (a matrix or data-frame) is computed with respect to the chosen reference data set. The rows of the matrices correspond to genes, and the columns to experimental conditions. The user can decide wether plotting a graph to enhance the deepest point against the rest of samples, or to show the MBD values according to some color palette, representing the genes in parallel coordinates. Alternatively, given percentages of most central curves can be used to display bands of curves, instead of the individual curves. The appearance of the plot can be adjusted with the Graphical options button. In addition, the outputs to be stored, i.e., the depth value of each data point and its position from centre outwards, can be selected with the corresponding button.

Author(s)

Sara Lopez-Pintado <sl2929@columbia.edu> and

Aurora Torrente <etorrent@est-econ.uc3m.es>

---

computeScaleCurve   *Computation and Representation of the Scale Curve*

Description

computeScaleCurve computes the scale curve of a given group, based on the MBD, at a given value p as the area of the band delimited by the [np] most central observations, where [np] is the largest integer smaller than np.

Details

The scale curve measures the increase in the area of the band determined by the fraction p most central curves, where p moves from 0 to 1, thus providing a measure of the sample dispersion. If the data set is represented in parallel coordinates, then the area is computed using the trapezoid formula.

Value

computeScaleCurve plots the scale curve of the active data set, which can contain a single group or several. In the latter case, a vector of labels can be provided to compute the scale curve for each group. The Y-coordinates used in the plot, i.e., the scale curve values at each point p can be stored as a vector, if only one group is present, or as a list, if there are several groups.

Author(s)

Sara Lopez-Pintado <sl2929@columbia.edu> and

Aurora Torrente <etorrent@est-econ.uc3m.es>

References

Lopez-Pintado, S. *et al.* (2010). Robust depth-based tools for the analysis of gene expression data. *Biostatistics*, 11 (2), 254-264.

See Also

computeMBD

---

computeTmean   *Modified Band Depth-Based Alpha Trimmed Mean*

Description

computeTmean computes the mean of the deepest observations within the sample, their depths given by the Modified Band Depth, trimming out the proportion alpha of the outest observations.

Details

The rows of active data set, corresponding to genes, are ordered from center outward, that is, starting with the deepest one(s) and ending with the less deep one(s), according to MBD. The alpha-trimmed mean is computed by first removing the proportion alpha of less deep points, and then computing the component-wise average of the remaining observations. The user can select the proportion of external points that are trimmed out and decide whether plotting the data set in parallel coordinates along with the trimmed mean. In addition, the plot can show the usual mean and the data points used to compute the trimmed mean.

Author(s)

Sara Lopez-Pintado <sl2929@columbia.edu> and

Aurora Torrente <etorrent@est-econ.uc3m.es>

See Also

computeMBD

---

plotCentralCurves   *Plot of the p Percent Most Central Curves of a Dataset*

Description

The function plotCentralCurves distinctly plots the p

Details

The rows of active data set, corresponding to genes, are ordered from center outwards, according to MBD. Then the [np/100] most central observations, where [x] is the largest integer smaller than x, and the remaining most external ones are plotted distinctly. The user can select the proportion of central curves points that are enhanced, and also assign a color palette to the most central ones to facilitate the understanding of the data structure.

Author(s)

Sara Lopez-Pintado <sl2929@columbia.edu> and Aurora Torrente <etorrent@est-econ.uc3m.es>

See Also

computeMBD, computeTmean, computeScaleCurve

Examples

```
data(prostate, package="depthTools")
prostate <- as.data.frame(prostate)
centralPlot(prostate, p = 0.5, col.c = '#ff0000',  col.e = '#C0C0C0' , lty=c(1,3) ,
    gradient = FALSE, gradient.ramp = c('#ff0000', '#ffd700'))
```

---

runDS   *Distance to the Trimmed Mean Classification Method*

Description

Implementation of the classification technique based on assigning each observation to the group that minimizes the distance of the observation to the trimmed mean of the group.

Details

The user can choose the learning and test sets, as well as the labels corresponding to the learning set. The DS method proceeds by first computing the alpha trimmed mean corresponding to each group from the learning set, then computing the distance from a new observation to each trimmed mean. The new sample will then be assigned to the group that minimizes such distance. At the moment, only the Euclidean distance is implemented. The predicted labels can be stored as a vector.

Author(s)

Sara Lopez-Pintado <sl2929@columbia.edu> and

Aurora Torrente <etorrent@est-econ.uc3m.es>

References

Lopez-Pintado, S. *et al.* (2010). Robust depth-based tools for the analysis of gene expression data. *Biostatistics*, 11 (2), 254-264.

See Also

computeTmean, runTAD

---

runRtest   *Rank Test Based on the Modified Band Depth*

Description

runRtest performs the rank test based on the MBD to decide whether two samples come from a single parent distribution.

Details

Given a population P from which a sample of n vectors is drawn, and another population P' from which a second sample of m vectors is obtained, assume there is a third reference sample (from the same population as the largest sample), whose size is also larger than n and m. The user selects the data sets containing the samples from both populations to be tested and the number of elements n and m to be included in each sample. runRtest identifies the largest sample as the one to be split into test and reference samples and verifies whether there are enough observations to run the test. Then, the proportions R and R' of elements from the reference sample whose depths are less or equal than those from the other samples, relative to the reference one, respectively, are computed and ordered from smallest to highest, giving them a rank from 1 to n+m. The statistic sum of the ranks of values R' (from the second population) has the distribution of a sum of m elements randomly drawn from 1 to n+m without replacement. The output is a list containing the p-value of the rank test and the test statistic value.

Author(s)

Sara Lopez-Pintado <sl2929@columbia.edu> and

Aurora Torrente <etorrent@est-econ.uc3m.es>

References

Lopez-Pintado, S. *et al.* (2010). Robust depth-based tools for the analysis of gene expression data. *Biostatistics*, 11 (2), 254-264.

See Also

computeTmean

---

runTAD   *Weighted Trimmed Mean Distance Classification Method*

Description

Implementation of the classification technique based on assigning each observation to the group that minimizes the trimmed average distance of the given observation to the deepest points of each group in the learning set, weighted by the depth of these points in their own group.

Details

The user can choose the learning and test sets, as well as the labels corresponding to the learning set. The TAD method classifies a given observation x into one of g groups, of sizes `n1,...,ng`, but taking into account only the `m=min{n1,...,ng}` deepest elements of each group in the learning set. Additionally, this number can be reduced in a proportion `alpha`. The distance of x to these m elements is averaged and weighted with the depth of each element with respect to its own group. The predicted labels can be stored as a vector.

Author(s)

Sara Lopez-Pintado `<sl2929@columbia.edu>` and

Aurora Torrente `<etorrent@est-econ.uc3m.es>`

References

Lopez-Pintado, S. *et al.* (2010). Robust depth-based tools for the analysis of gene expression data. *Biostatistics*, 11 (2), 254-264.

See Also

computeTmean, runDS

Index