Package 'ABCstats'

April 1, 2022

Type Package

Title Adaptive Box-Cox Transformation for Improving Data Normality

Version 1.0.0

Author Huaxu Yu

Maintainer <hxyu@chem.ubc.ca>

Description Feature-specific data transformation to improve data normality. The power
parameter lambda in Box-Cox transformation is optimized using Shapiro-Wilk normality test
of the input data.

License MIT + file LICENSE

Encoding UTF-8

LazyData true

Depends R (>= 4.0.0)

RoxygenNote 7.1.2

URL https://github.com/Waddlessss/ABCstats

Imports VIM, mice

R topics documented:

ABCtransform   *Adaptive Box-Cox transformation*

Description

Feature-specific data transformation to improve data normality in untargeted metabolomics.

Usage

```
ABCtransform(
  FeatureTable,
  Impt = "default",
  LambdaRange = c(-3, 3),
  GapIdentifier = 0,
  SampleInCol = TRUE,
  Output = FALSE
)
```

Arguments

FeatureTable	Data frame with features in row and samples in column (default).
Impt	A single string specifying the imputation method to be used.
LambdaRange	A numeric vector indicating the optimization range of lambda value.
GapIdentifier	A single value to identify gaps in `FeatureTable`.
SampleInCol	`TRUE` if samples are in column. `FALSE` if samples are in row.
Output	`TRUE` will output the result table in the current working directory.

Details

Adaptive Box-Cox (ABC) transformation is designed to improve the data normality in untargeted metabolomics. ABC transformation contains three modules:
1. Data pre-treatment by gap-filling and data scaling
2. Optimize lambda value for ABC transformation
3. Perform the ABC data transformation using the optimized lambda value

`FeatureTable` contains measured signal intensities of metabolic features, with features in row and samples in column (default). The column names should be sample names, and the first row should be sample group names (e.g. control, case).

Four imputation methods are provided here:
1. `default`, default imputation method by ABC transformation
2. knn, KNN method supported by **VIM** package. See [kNN](#) for details.
3. `addition`, only replace gaps with identical small values (min/5).
4. **rf**, random forest method supported by **mice** package. See [mice](#) for details. This method is not recommended due to long calculation time when feature number > 200.

Value

This function returns the transformed data frame.

References

To be updated.

Examples

```
TransformedTable = ABCtransform(DemoData)
```

Imputation	*Data imputation*

Description

Missing value imputation by different methods.

Usage

```
Imputation(
  FeatureTable,
  Impt = "default",
  GapIdentifier = 0,
  SampleInCol = TRUE,
  Output = FALSE
)
```

Arguments

FeatureTable	Data frame with features in row and samples in column (default).
Impt	A single string specifying the imputation method to be used.
GapIdentifier	A numeric vector indicating the optimization range of lambda value.
SampleInCol	TRUE if samples are in column. FALSE if samples are in row.
Output	TRUE will output the result table in the current working directory.

Details

Four imputation methods are provided here:
1. default, default imputation method by ABC transformation
2. knn, KNN method supported by **VIM** package. See kNN for details.
3. addition, only replace gaps with identical small values (min/5).
4. **rf**, random forest method supported by **mice** package. See mice for details. This method is not recommended due to long calculation time when feature number > 200.

FeatureTable contains measured signal intensities of metabolic features, with features in row and samples in column (default). The column names should be sample names, and the first row should be sample group names (e.g. control, case).

Value

This function returns the imputed data frame.

References

To be updated.

Examples

```
ImputedTable = Imputation(DemoData)
```

Index