Package 'trendsegmentR'

May 3, 2019

Title Linear Trend Segmentation and Point Anomaly Detection

Version 1.0.0

Description Performs the detection of point anomalies and linear trend changes
for univariate time series by implementing the bottom-up unbalanced wavelet transformation
proposed by H. Maeng and P. Fryzlewicz (2019) <http://personal.lse.ac.uk/maengh/>.
The estimated number and locations of the change-points are returned with the
piecewise-linear estimator for signal.

Depends R (>= 3.6.0)

License GPL

Encoding UTF-8

RoxygenNote 6.1.1

LazyData true

NeedsCompilation no

Maintainer Hyeyoung Maeng <h.maeng@lse.ac.uk>

Repository CRAN

Author Hyeyoung Maeng [aut, cre],
Piotr Fryzlewicz [aut]

Date/Publication 2019-05-03 15:20:03 UTC

R topics documented:

filter 2
invTGUW 3
orthmatrix 4
TGUW 5
thresholding 6
trendsegment 7
trendsegmentR 9

Index **10**

filter *Finding a detail filter vector for computing a detail coefficient*

Description

This function is used inside TGUW function but is typically not called directly by the user. This returns a vector of detail filter, a weight vector of the triplet of smooth coefficients where the corresponding detail coefficient is obtained as a weighted sum of three smooth coefficients.

Usage

```
filter(a)
```

Arguments

a	A vector with length 6. First three elements contain a triplet selected from the weight vector of constancy and the last three elements correspond to a selected triplet of the weight vector of linearity.

Details

The detail filter is obtained in such a way to produce zero detail coefficient only when the corresponding (raw) observations in merging regions have a perfect linear trend, as the detail coefficient itself represents the extent of non-linearity in the corresponding region of data. This implies that the smaller the size of the detail coefficient, the closer the alignment of the corresponding data section with linearity. For details, see H. Maeng and P. Fryzlewicz (2019), Detecting linear trend changes and point anomalies in data sequences, preprint.

Value

df	The detail filter vector of length 3 which is used as a weight vector for the corresponding triplet of smooth coefficients.

Author(s)

Hyeyoung Maeng, <h.maeng@lse.ac.uk>

See Also

TGUW

Examples

```
x <- c(rep(1, 3), 1:3)
filter(x)

y <- c(rep(1, 3), 4:6)
filter(y) # same as filter(x)
```

`invTGUW` *Inverse Tail-Greedy Unbalanced Wavelet (TGUW) transformation*

Description

This function performs the inverse TGUW transformation by undoing the orthonormal transformation of `TGUW` in reverse order. Details of the inverse TGUW transformation can be found in H. Maeng and P. Fryzlewicz (2019), Detecting linear trend changes and point anomalies in data sequences, preprint.

Usage

```
invTGUW(ts.obj)
```

Arguments

`ts.obj`	A list returned by `thresholding`.

Value

`ts.obj`	The modified ts.obj is the result of the inverse TGUW transformation and `ts.coeffs` now presents the estimated piecewise-linear signal of the data.

Author(s)

Hyeyoung Maeng, <h.maeng@lse.ac.uk>

See Also

`trendsegment`, `TGUW`, `thresholding`

Examples

```
x <- c(1:10, 0, rep(5,9))
n <- length(x)
x <- x + rnorm(n)
tguwfit <- TGUW(x)
th.const <- 1.3
lambda <- (stats::mad(diff(diff(x)))/sqrt(6)) * sqrt(2 * log(n)) * th.const
thrfit <- thresholding(ts.obj = tguwfit, lambda = lambda)
invfit <- invTGUW(ts.obj = thrfit)
invfit
```

orthmatrix *Finding an orthonormal matrix based on a detail filter vector*

Description

This function is used inside TGUW function but is typically not called directly by the user. This gives an orthonormal matrix with dimension 3 by 3 by computing two low filter vectors based on a given detail filter vector. The returned orthonormal matrix is firstly used in the orthonormal transformation of TGUW when updating three neighbouring smooth coefficients into one detail and two smooth coefficients, and its inverse matrix is used in the inverse TGUW transformation (invTGUW).

Usage

```
orthmatrix(d)
```

Arguments

d	A detail filter returned by filter which has a form of a vector with length 3.

Details

For details, see H. Maeng and P. Fryzlewicz (2019), Detecting linear trend changes and point anomalies in data sequences, preprint.

Value

M	The orthonormal matrix with dimension 3 by 3 which is used in TGUW and invTGUW.

Author(s)

Hyeyoung Maeng, <h.maeng@lse.ac.uk>

See Also

TGUW, invTGUW, filter

Examples

```
x <- c(rep(1, 3), 1:3)
df <- filter(x) # detail filter
orthmatrix(df)
```

TGUW *Tail-Greedy Unbalanced Wavelet (TGUW) transformation of a vector*

Description

Performs the bottom-up unbalanced wavelet decomposition. Details of the TGUW transformation can be found in H. Maeng and P. Fryzlewicz (2019), Detecting linear trend changes and point anomalies in data sequences, preprint.

Usage

```
TGUW(x, p = 0.01)
```

Arguments

x	An input vector to be decomposed.
p	Proportion of all possible remaining merges which specifies the number of merges allowed in a single pass over the data. The default is 0.01.

Value

A list with the followings:

x	The original input vector x.
n	The length of x.
twotogether	A vector indicating locations of the detail coefficients returned by Type 3 merges (merging two sets of paired smooth coefficients). This is used in thresholding to apply the "two together" rule which makes both detail coefficients (paired by a Type 3 merge) survived if at least one of their size is over threshold.
merging.hist	An array of dimension 4 by 3 by n-2 which has the full record of the n-2 merges in the TGUW transformation. Each matrix contains the information of each merge. The first row shows the indices of merged smooth coefficients in increasing order and the second row gives the value of detail filter coefficients which is the weight vector for computing the corresponding detail coefficient. The third row shows the (detail coefficient, first smooth coefficient, second smooth coefficient) obtained by an orthonormal transform. The fourth row gives the balancedness of merging. If it is Type 1 merging (three initial smooth coefficients) then the fourth row is always (1/3, 1/3, 1/3). In Type 2 and Type 3 mergings, the values depend on the ratio of the length of the left and right wings to the entire merged region and only first two components of the fourth row are filled with the corresponding ratios (sum to 1) but the third one is left as NA.
ts.coeffs	The transformed x by the TGUW transformation.

Author(s)

Hyeyoung Maeng, <h.maeng@lse.ac.uk>

See Also

trendsegment, thresholding, invTGUW

Examples

```
x <- c(1:10, 0, rep(5,9))
n <- length(x)
x <- x + rnorm(n)
tguwfit <- TGUW(x)
tguwfit
```

`thresholding` *Noise reduction from the sequence of detail coefficients returned by the Tail-Greedy Unbalanced Wavelet (TGUW) transformation*

Description

This function performs the thresholding of the detail coefficients returned by the Tail-Greedy Unbalanced Wavelet (TGUW) transformation. The denoising is achieved by a prespecified threshold in a "connected" way in that it prunes the branches if and only if the detail coefficient itself and all of its children coefficients are below some thresholds in its size. Also, the "two together" rule is applied to any paired detail coefficients returned by Type 3 merging (merging two sets of paired smooth coefficients) in the sense that both detail coefficients should be survived if at least one of their size is over threshold. For details, see H. Maeng and P. Fryzlewicz (2019), Detecting linear trend changes and point anomalies in data sequences, preprint.

Usage

```
thresholding(ts.obj, lambda, bal = 0)
```

Arguments

`ts.obj`	A list returned by `TGUW`.
`lambda`	The magnitude of the threshold. It has a form of `sigma * th.const * sqrt(2 * log(n))` where `n` is the length of input data `x`, the default of `th.const` is 1.3 and the `sigma` can be estimated by Median Absolute Deviation (MAD) method under the Gaussian assumption for noise.
`bal`	The minimum ratio of the length of the shorter region to the length of the entire merging region especially when the merges of Type 2 (merging one initial and a paired smooth coefficient) or of Type 3 (merging two sets of (paired) smooth coefficients) are performed. Only triplets which satisfy this balancedness condition survives in denoising. Point anomalies can be detected only if `bal < 1/n`. The default is set to 0 for detecting both point anomalies and linear trend changes.

Value

`ts.obj`	The modified ts.obj containing zero detail coefficients in the `merging.hist` if not survived from thresholding.

Author(s)

Hyeyoung Maeng, <h.maeng@lse.ac.uk>

See Also

trendsegment, TGUW, invTGUW

Examples

```
x <- c(1:10, 0, rep(5,9))
n <- length(x)
x <- x + rnorm(n)
tguwfit <- TGUW(x)
th.const <- 1.3
lambda <- (stats::mad(diff(diff(x)))/sqrt(6)) * sqrt(2 * log(n)) * th.const
thrfit <- thresholding(ts.obj = tguwfit, lambda = lambda)
thrfit
```

---

trendsegment *Detecting linear trend changes and point anomalies for univariate time series*

Description

The main function of the package trendsegmentR. This function estimates the number and locations of change-points in linear trend of noisy data. The estimated change-points may contain point anomalies if any. It also returns the estimated signal, the best linear fit for each segment between a pair of adjacent change-points. The algorithm includes three steps, Tail-Greedy Unbalanced Wavelet (TGUW) transform (TGUW), thresholding (thresholding) and inverse TGUW transform (invTGUW).

Usage

```
trendsegment(x, th.const = 1.3, p = 0.01, bal = 0)
```

Arguments

x	A data vector to be examined for change-point detection.
th.const	Thresholding parameter used in thresholding. The default is 1.3 and the exact magnitude of the threshold is `sigma * th.const * sqrt(2 * log(n))` where `n` is the length of data sequence `x` and `sigma` is estimated by Median Absolute Deviation (MAD) method under the Gaussian assumption for noise.
p	Proportion of all possible remaining merges which specifies the number of merges allowed in a single pass over the data. This is used in TGUW and the default is 0.01.

`bal` The minimum ratio of the length of the shorter region to the length of the entire merging region especially when the merges of Type 2 (merging one initial and a paired smooth coefficient) or of Type 3 (merging two sets of (paired) smooth coefficients) are performed. If `bal < 1/n`, point anomalies can be detected, otherwise, the detection of point anomalies is not guaranteed. The default is set to 0 for detecting both point anomalies and linear trend changes.

Details

The algorithm is described in H. Maeng and P. Fryzlewicz (2019), Detecting linear trend changes and point anomalies in data sequences, preprint.

Value

A list with the following.

`x` The original input vector `x`.

`est` The estimated piecewise-linear signal of `x`.

`no.of.cpt` The estimated number of change-points.

`cpt` The estimated locations of change-points.

Author(s)

Hyeyoung Maeng, <h.maeng@lse.ac.uk>

See Also

`TGUW`, `thresholding`, `invTGUW`

Examples

```
x <- c(rep(0,100), -4, seq(0, 4, length.out = 100), rep(3, 100), seq(3, -1, length.out=99))
n <- length(x)
x <- x + rnorm(n)
tsfit <- trendsegment(x = x, bal = 0)
tsfit

plot(x, type = "b", ylim = range(x, tsfit$est))
lines(tsfit$est, col=2, lwd=2)
```

trendsegmentR	*trendsegmentR: A device for detecting multiple change-points corresponding to linear trend changes or point anomalies in one dimensional data*

Description

Performs the detection of point anomalies and linear trend changes by transforming the data through an adaptively constructed unbalanced wavelet basis in a hierarchical way. To start with, see the function `trendsegment`.

Author(s)

Hyeyoung Maeng, <h.maeng@lse.ac.uk>

References

H. Maeng and P. Fryzlewicz (2019), Detecting linear trend changes and point anomalies in data sequences, preprint (http://personal.lse.ac.uk/maengh/).

See Also

`trendsegment`, `TGUW`, `thresholding`, `invTGUW`

Index

filter, 2, *4*

invTGUW, 3, *4*, *6–9*

orthmatrix, 4

TGUW, *2–4*, 5, *7–9*
thresholding, *3*, *5*, *6*, 6, *7–9*
trendsegment, *3*, *6*, *7*, 7, *9*
trendsegmentR, *7*, 9
trendsegmentR-package *(*trendsegmentR*)*, 9