

Package ‘matrixProfile’

August 17, 2018

Type Package

Title Matrix Profile

Version 0.5.0

Date 2018-08-14

Author Donghwan Kim

Maintainer Donghwan Kim <donhkim9714@korea.ac.kr>

Description A simple and the early stage package for matrix profile based on the paper of Chia Michael Yeh, Yan Zhu, Liudmila Ulanova, Nurjahan Begum, Yifei Ding, Hoang Anh Dau, Diego Furtado Silva, Abdullah Mueen, and Eamonn Keogh (2016) <DOI:10.1109/ICDM.2016.0179>. This package calculates all-pairs-similarity for a given window size for time series data.

License GPL-3

Encoding UTF-8

LazyData yes

Repository CRAN

URL <https://github.com/ainsuotain/matrixprofile>

Depends R (>= 3.0.0), graphics, stats, TTR, zoo, fftw, signal

NeedsCompilation no

Date/Publication 2018-08-17 09:20:07 UTC

R topics documented:

mass	2
movmean	3
movstd	4
stamp	5
std	6
Index	8

`mass`*Mueen's ultra-fast Algorithm for Similarity Search (MASS)*

Description

Calculates a sliding dot products of given data.

Usage

```
mass(q, t)
```

Arguments

<code>q</code>	A query data for dot product.
<code>t</code>	A timeseries data for analysis.

Value

Dot products between query and all subsequences in Timeseries. In the paper, we can implicitly construct a distance matrix with this output values that are the result of these dot products.

Author(s)

Donghwan Kim
<ainsuotain@hanmail.net> <donhkim9714@korea.ac.kr> <dhkim2@bistel.com>

References

Yeh, C. C. M., Zhu, Y., Ulanova, L., Begum, N., Ding, Y., Dau, H. A., ... & Keogh, E. (2016) <DOI:10.1109/ICDM.2016.0179>. *2016 IEEE 16th International Conference on Data Mining (ICDM)*, Barcelona, 2016, pp. 1317-1322.
https://www.cs.unm.edu/~mueen/MASS_V2.m

Examples

```
dt = AirPassengers
dt = as.vector(dt)
par(mfrow = c(2,1))
plot(dt, type = "l")
dm <- mass(q = dt[1:10], t = dt[-c(1:10)])
plot(dm, tyep = "l")
```

movmean	<i>Moving mean</i>
---------	--------------------

Description

Calculates moving mean of input data.

Usage

```
movmean(x = x, w = w)
```

Arguments

x	A given input data.
w	A sliding window of length w.

Value

An array of local w-point mean values, where each mean is calculated over a sliding window of length w across neighboring elements of x. The window size is automatically truncated at the endpoints when there are not enough elements to fill the window. When the window is truncated, the average is taken over only the elements that fill the window. Output is the same size as x.

Author(s)

Donghwan Kim
<ainsuotain@hanmail.net> <dohkim9714@korea.ac.kr> <dhkim2@bistel.com>

References

<https://kr.mathworks.com/help/matlab/ref/movmean.html?lang=en>

Examples

```
x <- 1:10  
movmean(x, 3)
```

movstd

Moving std

Description

Calculates moving sample standard deviation of input data.

Usage

```
movstd(x = x, w = w)
```

Arguments

x	A given input data.
w	A sliding window of length w.

Value

An array of local w-point sample standard deviation values, where each sample standard deviation is calculated over a sliding window of length w across neighboring elements of x. The window size is automatically truncated at the endpoints when there are not enough elements to fill the window. When the window is truncated, the standard deviation is taken over only the elements that fill the window. Output is the same size as x.

Author(s)

Donghwan Kim
<ainsuotain@hanmail.net> <donhkim9714@korea.ac.kr> <dhkim2@bistel.com>

References

<https://kr.mathworks.com/help/matlab/ref/movstd.html?lang=en>

Examples

```
x <- 1:10  
movstd(x, 3)
```

stamp

Scalable Time series Anytime Matrix Profile (stamp)

Description

Calculates a matrix profile of given data using STAMP algorithm.

Usage

```
stamp(q, t, by = 10, isPlot = FALSE)
```

Arguments

q	A query data for dot product.
t	A timeseries data for analysis.
by	A parameter that indicates the progress of the process in the process of calculating the matrix profile. For example, if by is 10, the text is displayed in the console window every 10 percent.
isPlot	A parameter that determines whether or not to draw a plot in the middle of calculating a matrix profile. The default value is FALSE.

Details

The matrix profile is calculated by the self join method using the STAMP algorithm. One of the key features of the STAMP algorithm is the attribute anytime. In other words, because the matrix profile is computed rather than randomly, the computation speed is the same, but it is quickly optimized.

Value

An object of class `stamp.models`.

MP	A matrix profile computed by given data.
MPI	A matrix profile index computed by given data.
MTI	A motif index of matrix profile. Unlike in the original paper, it denotes the pair of motif index with the smallest value of matrix profile.

Note

This package is an early version and will be updated in the neae future. Also note that it is very slow for data with more than 10,000 data points. Since it is not optimized basic functions(e.g. `movmean`, `movstd`) for computation and is due to R's own limitations.

Author(s)

Donghwan Kim
<ainsuotain@hanmail.net> <donhkim9714@korea.ac.kr> <dhkim2@bistel.com>

References

Yeh, C. C. M., Zhu, Y., Ulanova, L., Begum, N., Ding, Y., Dau, H. A., ... & Keogh, E. (2016) <DOI:10.1109/ICDM.2016.0179>. *2016 IEEE 16th International Conference on Data Mining (ICDM)*, Barcelona, 2016, pp. 1317-1322.
<http://www.cs.ucr.edu/~eamonn/MatrixProfile.html>

See Also

[mass](#) (in package **matrixProfile**)

Examples

```
# data input
dt = AirPassengers
dt = as.vector(dt)

# generates matrix profile
stamp <- stamp(q = dt[1:12], t = dt[-c(1:12)])

# plotting
par(mfrow = c(2,1))
plot(dt, type = "l", main = "Original Timeseries")
plot(stamp$MP, type = "l", main = "Matrix Profile", xlim = c(0, length(dt)))
```

std	<i>Sample standard deviation</i>
-----	----------------------------------

Description

Calculates sample standard deviation of input data.

Usage

```
std(x)
```

Arguments

x A given input data.

Details

This function is slightly different from the **base** function `sd`.

Value

An sample standard deviation of given data.

Author(s)

Donghwan Kim
<ainsuotain@hanmail.net> <donhkim9714@korea.ac.kr> <dhkim2@bistel.com>

References

https://en.wikipedia.org/wiki/Standard_deviation

Examples

```
x <- 1:10
sd(x) # for comparison
std(x) # see difference
```

Index

- *Topic **distance profile**
 - stamp, [5](#)
- *Topic **matrix profile**
 - stamp, [5](#)
- *Topic **time series motif discovery**
 - stamp, [5](#)

- mass, [2](#), [6](#)
- movmean, [3](#)
- movstd, [4](#)

- stamp, [5](#)
- std, [6](#)