Package ‘AnomalyScore’

November 21, 2024

Title Anomaly Scoring for Multivariate Time Series

Version 0.1

Description Compute an anomaly score for multivariate time series based on the k-nearest neighbors algorithm. Different computations of distances between time series are provided.

License GPL (>= 3)

Imports dtw, class, astsa, transport, marima, TSA, RANN, MASS, mvLSW, stats

Encoding UTF-8

RoxygenNote 7.3.1

NeedsCompilation no

Author Guillermo Granados [aut, cre] (<https://orcid.org/0000-0003-4868-6163>)

Maintainer Guillermo Granados <guillermo.granadosgarcia@outlook.com>

Repository CRAN

Date/Publication 2024-11-21 08:00:13 UTC

Contents

`all_bands` *Pairwise band generation in a multivariate time series*

Description

takes all the pairs of series in a multivariate set and compute the band between each pair od series

Usage

```
all_bands(series)
```

Arguments

`series`	A matrix of n columns representing a multivariate time series, each column is a univariate time series

Value

a list with two elements the lowerbounds of all (n)(n-1)/2 pairs and the upperbounds of the pairwise bands

See Also

Band Depth Clustering for Nonstationary Time Series and Wind Speed Behavior (2018) Tupper et al

Examples

```
X=matrix( rnorm(200), ncol=10  )
all_bands(X)
```

`Anomalyscoresframe` *Anomaly score computation for a set of distances*

Description

Computes anomaly scores for a selection of different distances for a single dataset.

Usage

```
Anomalyscoresframe(unit, knn, measures, dparams)
```

Arguments

`unit`	A matrix representing a multivariate time series where each column is a univariate time series.
`knn`	number of nearest neighbors to consider for the anomaly scores
`measures`	vector with the indexes of the selected measures 1=Cort, 2=Wasserstein, 3=Mahalanobis, 4=Normalized Cort, 5=Coherence, 6=PDC, 7=CGCI,8=RGPDC, 9=PMIME, 10=mvLWS, 11=Band depth
`dparams`	a list where each element is a list with all the parameters necessary to compute the selected distances. If the distance does not need further parameters then define an empty list

Value

A dataframe with the names of series in unit as a column called "series" and the corresponding scores computed for each distance. The rank is ordered with respect to the first measure in the measures index vector

See Also

Guillermo Granados, and Idris Eckley. "Electricity Demand of Buildings Benchmarked via Regression Trees on Nearest Neighbors Anomaly Scores"

Examples

```
unit=matrix( rnorm(500), ncol=5  )
measures= c(1,5,11 ) # Cort, Coherence and Band depth
knn=3
dparams=list(
  list(k=2),
  list( span1=2, span2=2, period = 5),
  list( )
)
Anomalyscoresframe(unit, knn,measures, dparams)
```

`Cort` *Temporal correlation coefficient*

Description

Return the temporal correlations of two time series by first taking the lag one differences of each series and the computing the correlation coefficient.

Usage

```
Cort(S1, S2)
```

Arguments

`S1`	A vector representing a univariate time series
`S2`	A second vector representing a univariate time series

Value

A coefficient in the interval $[-1, 1]$ representing the lag 1 correlation

See Also

Douzal-Chouakria, Ahlame, and Cecile Amblard. "Classification Trees for Time Series." Pattern Recognition 45, no. 3 (March 2012): 1076-91. doi:10.1016/j.patcog.2011.08.018

Examples

```
S1=rnorm(100)
S2=rnorm(100)
Cort(S1, S2)
```

`DEcort` *Distance based on value and behavior of the time series*

Description

Return a weighted distance based on a weighted sum of the Euclidean norm and the temporal correlation coefficient. The distance is inflated in the presence of NA compensating for the lack of information.

Usage

```
DEcort(k, S1, S2)
```

Arguments

`k`	The parameter \$k\$ controls the contribution of the sum of squares comparison as a value-based metric and the \$Cort\$ quantity as a behavioral metric; when \$k=0\$, then the distance is equal to the value-based metric, on the other hand, when \$k=6\$ the distance is mainly determined by the value of the temporal correlation \$Cort\$.
`S1`	A vector representing a univariate time series
`S2`	A second vector representing a univariate time series

Value

a non-zero value

See Also

Douzal-Chouakria, Ahlame, and Cecile Amblard. "Classification Trees for Time Series." Pattern Recognition 45, no. 3 (March 2012): 1076-91. doi:10.1016/j.patcog.2011.08.018

Examples

```
S1=rnorm(100)
S2=rnorm(100)
k=1
DEcort(k,S1, S2)
```

DEcortNorm

Normalized version of the Cort distance the modification is based on using the coefficient of variation rather than euclidean distance, performed by normalizing by the absolute value of the total differences of the series.

Description

Normalized version of the Cort distance the modification is based on using the coefficient of variation rather than euclidean distance, performed by normalizing by the absolute value of the total differences of the series.

Usage

```
DEcortNorm(k, S1, S2)
```

Arguments

k	The parameter k controls the contribution of the sum of squares comparison as a value-based metric and the $Cort$ quantity as a behavioral metric; when $k=0$, then the distance is equal to the value-based metric, on the other hand, when $k=6$ the distance is mainly determined by the value of the temporal correlation $Cort$.
S1	A vector representing a univariate time series
S2	A second vector representing a univariate time series

Value

a non-zero value

See Also

Granados-Garcia, and Idris Eckley. "Building Electricity Demand Benchmarking via a Regression Trees on Anomaly Scores"

Examples

```
S1=rnorm(100)
S2=rnorm(100)
k=1
DEcortNorm(k,S1, S2)
```

---

distance_matrix_banddepth

Pairwise distance matrix based on the band depth distance

Description

Pairwise distance matrix based on the band depth distance

Usage

```
distance_matrix_banddepth(unit)
```

Arguments

unit	A matrix representing a multivariate time series where each column is a univariate time series.

Value

a matrix with pairwise distances

See Also

Band Depth Clustering for Nonstationary Time Series and Wind Speed Behavior (2018) Tupper et al

Examples

```
X=matrix( rnorm(2000), ncol=10  )
distance_matrix_banddepth(unit=X)
```

---

distance_matrix_CGCI *Pairwise distance matrix based on the conditional Granger causality index*

Description

Pairwise distance matrix of a multivariate time series based on the the conditional Granger causality index distance between two series

Usage

```
distance_matrix_CGCI(unit, pmax)
```

Arguments

unit	A matrix representing a multivariate time series where each column is a univariate time series.
pmax	maximum order(lag) of the VAR model to be considered

Value

a matrix with pairwise distances

See Also

Siggiridou, Elsa, and Dimitris Kugiumtzis. "Granger Causality in Multivariate Time Series Using a Time-Ordered Restricted Vector Autoregressive Model." IEEE Transactions on Signal Processing 64, no. 7 (April 2016): 1759-73. doi:10.1109/TSP.2015.2500893

Examples

```
X=matrix( rnorm(2000), ncol=10  )
pmax=4
distance_matrix_CGCI(unit=X, pmax)
```

distance_matrix_coherence

Distance matrix from a coherence measure

Description

Pairwise distance matrix of a multivariate time series based on computing the squared coherence and transformed it to represent a distance at a specific frequency

Usage

```
distance_matrix_coherence(unit, span1, span2, period)
```

Arguments

`unit`	A matrix representing a multivariate time series where each column is a univariate time series.
`span1`	Odd integer giving the widths of modified Daniell smoothers to be used to smooth the periodogram. Refers to the bandwidth of the smoothing process.
`span2`	Odd integer giving the widths of modified Daniell smoothers to be used to smooth the periodogram. Control another level of smoothing to the spectral density estimation without altering the peaks
`period`	Integer referencing the index of the frequency to use for the distance. It gives the Hertz or periods per unit of time; i.e., if the sampling is per minute, and each hour cycle is the period of interest

Value

a matrix with pairwise distances

See Also

For more details, check the `astsa` package documentation on CRAN or visit the GitHub repository https://github.com/nickpoison/astsa.

Examples

```
X=matrix( rnorm(2000), ncol=10  )
span1=2
span2=2
period=3
distance_matrix_coherence(unit=X, span1, span2, period )
```

distance_matrix_cort *Distance matrix from a pattern recognition distance*

Description

pairwise distance matrix of a multivariate time series based on a value (Euclidean distance) and behavior (temporal correlation) measures

Usage

```
distance_matrix_cort(k, unit)
```

Arguments

k	The parameter k controls the contribution of the sum of squares comparison as a value-based metric and the $Cort$ quantity as a behavioral metric; when $k=0$, then the distance is equal to the value-based metric, on the other hand, when $k=6$ the distance is mainly determined by the value of the temporal correlation $Cort$.
unit	A matrix representing a multivariate time series where each column is a univariate time series.

Value

a matrix with pairwise distances

See Also

Douzal-Chouakria, Ahlame, and Cecile Amblard. "Classification Trees for Time Series." Pattern Recognition 45, no. 3 (March 2012): 1076-91. doi:10.1016/j.patcog.2011.08.018

Examples

```
X=matrix( rnorm(200), ncol=10  )
k=2
distance_matrix_cort(k,X)
```

distance_matrix_cortNorm
Normalized distance matrix from a pattern recognition distance

Description

pairwise distance matrix of a multivariate time series based on a value (Coefficient of variation) and behavior (temporal correlation) measures

Usage

```
distance_matrix_cortNorm(k, unit)
```

Arguments

k	The parameter k controls the contribution of the sum of squares comparison as a value-based metric and the $Cort$ quantity as a behavioral metric; when $k=0$, then the distance is equal to the value-based metric, on the other hand, when $k=6$ the distance is mainly determined by the value of the temporal correlation $Cort$.
unit	A matrix representing a multivariate time series where each column is a univariate time series.

Value

a matrix with pairwise distances

See Also

Guillermo Granados, and Idris Eckley. "Electricity Demand of Buildings Benchmarked via Regression Trees on Nearest Neighbors Anomaly Scores"

Examples

```
X=matrix( rnorm(200), ncol=10  )
k=2
distance_matrix_cortNorm(k,X)
```

distance_matrix_dtw *Normalized distance matrix from dynamic time-warping distance*

Description

pairwise distance matrix of a multivariate time series based on a minimal mapping between two time series weighted by temporal correlation

Usage

```
distance_matrix_dtw(k, unit, maxwindow)
```

Arguments

`k`	The parameter k controls the contribution of the sum of squares comparison as a value-based metric and the $Cort$ quantity as a behavioral metric; when $k=0$, then the distance is equal to the value-based metric, on the other hand, when $k=6$ the distance is mainly determined by the value of the temporal correlation $Cort$.
`unit`	A matrix representing a multivariate time series where each column is a univariate time series.
`maxwindow`	the maximum shift allowed between time series points.

Value

a matrix with pairwise distances

See Also

For more details, check the `dtw` package documentation on CRAN

Examples

```
X=matrix( rnorm(200), ncol=10  )
k=2
maxwindow=10
distance_matrix_dtw(k,X,maxwindow)
```

distance_matrix_mahalanobis

Pairwise distance matrix based on the mahalanobis distance

Description

Pairwise distance matrix of a multivariate time series based on the Mahalanobis distance between two series, modified to consider the different scales of series

Usage

```
distance_matrix_mahalanobis(unit)
```

Arguments

unit	A matrix representing a multivariate time series where each column is a univariate time series.

Value

a matrix with pairwise distances

See Also

Prekopcsak, Zoltan, and Daniel Lemire. "Time Series Classification by Class-Specific Mahalanobis Distance Measures." Advances in Data Analysis and Classification 6, no. 3 (October 2012): 185-200. doi:10.1007/s1163401201106

Examples

```
X=matrix( rnorm(2000), ncol=10  )
distance_matrix_mahalanobis(unit=X )
```

distance_matrix_mvLWS

Pairwise distance matrix based on the multivariate locally wavelet partial coherence

Description

Pairwise distance matrix based on the multivariate locally wavelet partial coherence

Usage

```
distance_matrix_mvLWS(unit)
```

Arguments

`unit`	A matrix representing a multivariate time series where each column is a univariate time series.

Value

a matrix with pairwise distances

See Also

Park, Timothy, Idris A. Eckley, and Hernando C. Ombao. "Estimating Time-Evolving Partial Coherence Between Signals via Multivariate Locally Stationary Wavelet Processes." IEEE Transactions on Signal Processing 62, no. 20 (October 2014): 5240-50. doi:10.1109/TSP.2014.2343937

Examples

```
X=matrix( rnorm(2000), ncol=10  )
distance_matrix_mvLWS(unit=X)
```

`distance_matrix_PDC` *Distance matrix from a partial directed coherence measure (PDC)*

Description

Pairwise distance matrix of a multivariate time series based on the partial directed coherence among two series. The distance considers both directions of causality and transform it to give 0 in absence of causality between the series.

Usage

```
distance_matrix_PDC(unit, ar, period)
```

Arguments

`unit`	A matrix representing a multivariate time series where each column is a univariate time series.
`ar`	Integer vector containing all the lags considered for the vector autoregressive model
`period`	Integer referencing the index of the frequency to use for the distance. It gives the Hertz or periods per unit of time; i.e., if the sampling is per minute, and each hour cycle is the period of interest

Value

a matrix with pairwise distances

See Also

Guillermo Granados, and Idris Eckley. "Electricity Demand of Buildings Benchmarked via Regression Trees on Nearest Neighbors Anomaly Scores"

Examples

```
X=matrix( rnorm(2000), ncol=10  )
ar=c(1, 2)
period=10
distance_matrix_PDC(  unit=X, ar,  period )
```

`distance_matrix_PMIME` *Pairwise distance matrix based on the partial mutual information of mixed embedings (PMIME) method*

Description

Pairwise distance matrix based on the partial mutual information of mixed embedings (PMIME) method

Usage

```
distance_matrix_PMIME(unit, Lmax, Tl, nnei, A)
```

Arguments

`unit`	A matrix representing a multivariate time series where each column is a univariate time series.
`Lmax`	: the maximum delay to search for X and Y components for the mixed embedding vector ,default is 5.
`Tl`	: Tl steps ahead that the mixed embedding vector has to explain. Note that if Tl>1 the future vector is of length Tl and contains the samples at times t+1,..,t+Tl ,dafault is 1.
`nnei`	: number of nearest neighbors for density estimation ,default is 5
`A`	: the threshold for the ratio of CMI over MI of the lagged variables for the termination criterion.

Value

a matrix with pairwise distances

See Also

Kugiumtzis, D. "Direct-Coupling Information Measure from Nonuniform Embedding." Physical Review E 87, no. 6 (June 25, 2013): 062918. doi:10.1103/PhysRevE.87.062918

Examples

```
X=matrix( rnorm(300), ncol=3  )
Lmax=2
Tl=1
nnei=5
A=.95
distance_matrix_PMIME(unit=X, Lmax, Tl, nnei, A )
```

---

distance_matrix_RGPDC *Pairwise distance matrix based on the restricted generalized partial directed coherence*

Description

Pairwise distance matrix of a multivariate time series based on the the restricted generalized partial directed coherence distance between two series

Usage

```
distance_matrix_RGPDC(unit, pmax, period)
```

Arguments

unit	A matrix representing a multivariate time series where each column is a univariate time series.
pmax	maximum order(lag) of the VAR model to be considered
period	Integer referencing the index of the frequency to use for the distance. It gives the Hertz or periods per unit of time; i.e., if the sampling is per minute, and each hour cycle is the period of interest

Value

a matrix with pairwise distances

See Also

Siggiridou, Elsa, Vasilios K. Kimiskidis, and Dimitris Kugiumtzis. "Dimension Reduction of Frequency-Based Direct Granger Causality Measures on Short Time Series." Journal of Neuroscience Methods 289 (September 2017) : 64-74. doi:10.1016/j.jneumeth.2017.06.021

Examples

```
X=matrix( rnorm(2000), ncol=10  )
pmax=4
period=3
distance_matrix_RGPDC(unit=X, pmax, period)
```

distance_matrix_wasserstein *Distance matrix from based on the Wasserstein distance*

Description

Pairwise distance matrix of a multivariate time series based on the Wasserstein distance between the empirical distribution of the series

Usage

```
distance_matrix_wasserstein(unit)
```

Arguments

unit	A matrix representing a multivariate time series where each column is a univariate time series.

Value

a matrix with pairwise distances

See Also

For more details, check the `transport` package documentation on CRAN

Examples

```
X=matrix( rnorm(2000), ncol=10  )
distance_matrix_wasserstein(unit=X)
```

DTWcort *Extention of the dynamic time warping distance*

Description

This function uses the dtw() function from the dtw R package to compute a distance based on the mapping than minimizes the distance between two sets of points, the parameters chosen are the "Manhattan" distance to compute the differences between points and the "sakoechiba" window type. Important note: the dtw function does not accept NA values, therefore these types of values are removed.

Usage

```
DTWcort(k, S1, S2, maxwindow)
```

Arguments

k	The parameter \$k\$ controls the contribution of the sum of squares comparison as a value-based metric and the \$Cort\$ quantity as a behavioral metric; when \$k=0\$, then the distance is equal to the value-based metric, on the other hand, when \$k=6\$ the distance is mainly determined by the value of the temporal correlation \$Cort\$.
S1	A vector representing a univariate time series
S2	A second vector representing a univariate time series
maxwindow	the maximum shift allowed between time series points.

Value

A non-negative value representing the distance between two time series

See Also

Douzal-Chouakria, Ahlame, and Cecile Amblard. "Classification Trees for Time Series." Pattern Recognition 45, no. 3 (March 2012): 1076-91. doi:10.1016/j.patcog.2011.08.018

Examples

```
S1=rnorm(100)
S2=rnorm(100)
k=1
maxwindow=20
DTWcort(k,S1, S2,maxwindow)
```

---

dxy_bands *Band depth distance between 2 time series given a set of bands*

Description

Distance based on a depth concept, given a set of bands a modified Jaccard measure is compute between the sets of indices that two series share, the Jaccard distances then are averaged over all informative bands.

Usage

```
dxy_bands(allbands, x, y)
```

Arguments

allbands	a list with two elements the lowerbounds of all (n)(n-1)/2 pairs and the upper-bounds of the pairwise bands.
x	A vector representing a univariate time series
y	A vector representing a univariate time series

Value

A non-negative value representing the distance between two time series, based on the concept of band depth.

See Also

Band Depth Clustering for Nonstationary Time Series and Wind Speed Behavior (2018) Tupper et al

Examples

```
X=matrix( rnorm(200), ncol=10  )
M=all_bands(X)
dxy_bands(M,X[,1],X[,2] )
```

---

informative_bands *indexes where a series is within a specific band*

Description

Return the indicies in which the values of a series x are located within a band b, called the informative bands.

Usage

```
informative_bands(allbands, x)
```

Arguments

allbands	a list with two elements the lowerbounds of all (n)(n-1)/2 pairs and the upperbounds of the pairwise bands. Result of the function all_bands
x	A vector representing a univariate time series

Value

A vector with indices

See Also

Band Depth Clustering for Nonstationary Time Series and Wind Speed Behavior (2018) Tupper et al

Examples

```
X=matrix( rnorm(200), ncol=10  )
M=all_bands(X)
informative_bands(M,X[,1] )
```

kneighbors_distance_docall

K-Nearest neighbors algorithm to compute an anomaly score

Description

The method obtain a distance matrix and find the K-nearest neighbors of each series and sum their distances in the neighborhood. The sum is defined as the anomaly score, the series with higher scores implies their neighbors are far away and such a series is a potential outlier

Usage

```
kneighbors_distance_docall(knn, distance, dparams)
```

Arguments

`knn`	number of nearest neighbors to consider for the anomaly score
`distance`	function name of the available distance matrices
`dparams`	a list with all the parameters for the distance matrix

Value

A list of two elements with the anomaly scores and the distance matrix

See Also

Guillermo Granados, and Idris Eckley. "Electricity Demand of Buildings Benchmarked via Regression Trees on Nearest Neighbors Anomaly Scores"

Examples

```
X=matrix( rnorm(2000), ncol=10  )
distance=distance_matrix_coherence
dparams=list(unit=X, span1=2, span2=2, period = 5 )
knn=5
kneighbors_distance_docall(knn,distance, dparams)
```

`matrix_PDC` *Partial directed coherence matrix*

Description

Partial directed coherence matrix

Usage

```
matrix_PDC(unit, ar)
```

Arguments

unit	A Matrix containing the multivariate time series. Each column represents a univariate time series.
ar	Integer vector containing all the lags considered for the vector autoregressive model

Value

An real array of dimensions, ncol(unit), ncol(unit), n, where n is the number of frequencies at which the PDC is estimated.

Examples

```
X=matrix( rnorm(2000), ncol=10  )
ar=c(1, 2)
matrix_PDC(X, ar)
```

`mBTS` *modified Back-in-time Selection for vector AR parameters estimation*

Description

modified Back-in-time Selection for vector AR parameters estimation

Usage

```
mBTS(xM, responseindex, pmax)
```

Arguments

xM	the matrix of K time series (variables in columns)
responseindex	the index of the response variable in $\{1, \dots, K\}$
pmax	maximum order(lag) of the VAR model to be considered

Value

the matrix of all explanatory lagged variables in the DR model. The sequence of the lagged variables in 'lagM'

See Also

I. Vlachos and D. Kugiumtzis, "Backward-in-time selection of the order of dynamic regression prediction model," J. Forecast., vol. 32, pp. 685-701, 2013.

`mBTSCGCI` *computation of the conditional Granger causality index*

Description

computation of the conditional Granger causality index

Usage

```
mBTSCGCI(xM, responseindex, pmax)
```

Arguments

`xM`	the matrix of K time series (variables in columns)
`responseindex`	the index of the response variable in $\{1, \ldots, K\}$
`pmax`	maximum order(lag) of the VAR model to be considered

Value

the matrix of all the conditional Granger causality index across the series of a multivariate set.

See Also

Siggiridou, Elsa, and Dimitris Kugiumtzis. "Granger Causality in Multivariate Time Series Using a Time-Ordered Restricted Vector Autoregressive Model." IEEE Transactions on Signal Processing 64, no. 7 (April 2016): 1759-73. doi:10.1109/TSP.2015.2500893

mBTSRGPDC — *Restricted Generalized Partial Directed Coherence*

Description

partial directed coherence matrix values based on the mBTS algorithm for estimation of the VAR parameters

Usage

```
mBTSRGPDC(xM, pmax, freqs)
```

Arguments

`xM`	the matrix of K time series (variables in columns)
`pmax`	maximum order(lag) of the VAR model to be considered
`freqs`	frequencies at which the spectral density is estimated

Value

the matrix of all the Restricted Generalized Partial Directed Coherence index across the series of a multivariate set. #'

mBTS_Af_mat — *mBTS Vector Autoregressive coefficients fourier transform*

Description

DFT Vector Autoregressive coefficients matrix using the mBTS algorithm. This matrix is the base to compute the generalized partial directed coherence

Usage

```
mBTS_Af_mat(xM, responseindex, pmax, freqs)
```

Arguments

`xM`	the matrix of K time series (variables in columns)
`responseindex`	the index of the response variable in $\{1, \ldots, K\}$
`pmax`	maximum order(lag) of the VAR model to be considered
`freqs`	frequencies at which the spectral density is estimated

Value

the matrix of all the Restricted Generalized Partial Directed Coherence index across the series of a multivariate set.

`multilagmatrix` *multilagmatrix*

Description

multilagmatrix builds the set of explanatory variables for the dynamic regression model.

Usage

```
multilagmatrix(xM, responseindex, ordersV, indexV)
```

Arguments

`xM`	the matrix of K time series (variables in columns)
`responseindex`	the index of the response variable in $\{1, \ldots, K\}$
`ordersV`	vector of size 1xK of the maximum order for each of the K variables.
`indexV`	the vector of size 1 x K*pmax of zeros and ones e.g. if the component in position 2*pmax+3 is one, the third variable, lag 3, X3(t-3), is selected.

Value

the matrix of all explanatory lagged variables in the DR model. The sequence of the lagged variables in 'lagM'

See Also

Kugiumtzis, D. "Direct-Coupling Information Measure from Nonuniform Embedding." Physical Review E 87, no. 6 (June 25, 2013): 062918. doi:10.1103/PhysRevE.87.062918

`PMIME` *PMIME Partial mutual information from mixed embedding*

Description

computes the measure $R_{X->Y|Z}$ for all combinations of X and Y time series from the multivariate time series given in matrix 'allM', of size NxK, where Z contains the rest $K-2$ time series. The components of X,Y, and Z, are found from a mixed embedding aiming at explaining Y. The mixed embedding is formed by using the progressive embedding algorithm based on conditional mutual information (CMI). CMI is estimated by the method of nearest neighbors (Kraskov's method). The function is the same as PMIMEsig.m but defines the stopping criterion differently, using a fixed rather than adjusted threshold. Specifically, the algorithm terminates if the contribution of the selected lagged variable in explaining the future response state is small enough, as compared to a threshold 'A'. Concretely, the algorithm terminates if $I(x^F; w|wemb)/I(x^F; w, wemb) <= A$ where $I(x^F; w|wemb)$ is the CMI of the selected lagged variable w and the future response state x^F given the current mixed embedding vector, and $I(x^F; w, wemb)$ is the MI between x^F and the augmented mixed embedding vector $wemb, w$. We experienced that in rare cases the termination condition is not satisfied and the algorithm does not terminate. Therefore we included a second condition for termination of the algorithm when the ratio $I(x^F; w|wemb)/I(x^F; w, wemb)$ increases in the last two embedding cycles. The derived R measure indicates the information flow of time series X to time series Y conditioned on the rest time series in Z. The measure values are stored in a KxK matrix 'RM' and given to the output, where the value at position (i, j) indicates the effect from i to j (row to col), and the (i, i) components are zero. The library RANN was used for the nearest neighbor estimation of the mutual information

Usage

```
PMIME(allM, Lmax = 5, Tl = 1, nnei = 5, A = 0.03, showtxt = 1)
```

Arguments

`allM`	the N x K matrix of the K time series of length N.
`Lmax`	the maximum delay to search for X and Y components for the mixed embedding vector ,default is 5.
`Tl`	Tl steps ahead that the mixed embedding vector has to explain. Note that if T>1 the future vector is of length T and contains the samples at times t+1,..,t+T ,dafault is 1.
`nnei`	number of nearest neighbors for density estimation ,default is 5
`A`	the threshold for the ratio of CMI over MI of the lagged variables for the termination criterion.
`showtxt`	: if 0 or negative do not print out anything, if 1 print out the response variable index at each run, if 2 or larger print also info for each embedding cycle ,default is 1.

Value

RM: A K x K matrix containing the R values computed by PMIME using the surrogates for setting the stopping criterion. *ecC*: cell array of K components, where each component is a matrix of size E x 5, and E is the number of embedding cycles. For each embedding cycle the following 5 results are stored: 1. variable index, 2. lag index, 3. CMI of the selected lagged variable w and the future response state x^F given the current mixed embedding vector, I(x^F; w| wemb). 4. MI between x^F and the augmented mixed embedding vector wemb w, I(x^F; w,wemb). 5. The ration of 3. and 4.: I(x^F; w| wemb)/I(x^F; w,wemb)

See Also

Kugiumtzis, D. "Direct-Coupling Information Measure from Nonuniform Embedding." Physical Review E 87, no. 6 (June 25, 2013): 062918. doi:10.1103/PhysRevE.87.062918 sqnorms *Quadratic multiplication of a matrix M with respect to a matrix A: Conj(M) A M, where Conj() is the complex conjugate function*

Description

Quadratic multiplication of a matrix M with respect to a matrix A: Conj(M) A M, where Conj() is the complex conjugate function

Usage

```
sqnorms(M, A)
```

Arguments

M	A Matrix of dimension P by P
A	A squared Matrix of dimension P by P

Value

The squared root of the absolute values of the matrix result of the operation Conj(M) A M

Examples

```
M=matrix( rnorm(100), ncol=10  )
A=matrix( rnorm(100), ncol=10  )
sqnorms(M, A)
```

Index