

Package ‘HarmonicRegression’

April 1, 2015

Type Package

Title Harmonic Regression to One or more Time Series

Version 1.0

Date 2015-04-01

Author Paal O. Westermark

Maintainer Paal O. Westermark <pal-olof.westermark@charite.de>

Description Fits the first harmonics in a Fourier expansion to one or more time series. Trend elimination can be performed. Computed values include estimates of amplitudes and phases, as well as confidence intervals and p-values for the null hypothesis of Gaussian noise.

License MIT + file LICENSE

VignetteBuilder knitr

Suggests knitr

Depends R (>= 2.10)

NeedsCompilation no

Repository CRAN

Date/Publication 2015-04-01 19:27:09

R topics documented:

harmonic.regression	2
rna.nasc	3
rna.polya	4

Index	5
--------------	----------

harmonic.regression *Harmonic Regression*

Description

Estimates amplitudes and phases along with confidence intervals and p-values from a set of time series that may oscillate with a specified period. A model, per default

$$y = m + a\cos(\omega t) + b\sin(\omega t),$$

is fitted to the time series. This model is equivalent to the model

$$m + c\cos(\omega t - \phi),$$

with amplitude $c = \sqrt{a^2 + b^2}$ and phase $\phi = \text{atan2}(b, a)$. P-values for $c > 0$ (more precisely: either a or $b > 0$) are computed by an F-test. Confidence intervals for the amplitudes and phases are computed by a linear error propagation approximation.

Usage

```
harmonic.regression(inputts, inputtime, Tau = 24, normalize = TRUE,
  norm.pol = FALSE, norm.pol.degree = 1, trend.eliminate = FALSE,
  trend.degree = 1)
```

Arguments

inputts	Matrix of time series. Rows correspond to time points, columns to samples. If a vector is provided, it is coerced to a matrix.
inputtime	Vector of the time points corresponding to the row in the time series matrix.
Tau	Scalar giving the oscillation period to estimate and test for.
normalize	Boolean, set to TRUE if normalization is to be performed (default). Unless norm.pol=TRUE, normalization is performed by dividing with the mean.
norm.pol	Boolean, set to TRUE if a polynomial should be fitted to each time series, and used for normalization. In this case, each point in a time series will be divided by the value of the fitted polynomial. Defaults to FALSE.
norm.pol.degree	Scalar indicating the polynomial degree for the normalization (ignored if norm.pol=FALSE).
trend.eliminate	Boolean, set to TRUE if trend elimination is to be performed (see Details above). Defaults to FALSE.
trend.degree	Integer indicating the polynomial degree for the trend elimination, default is 1. Ignored when trend.eliminate=FALSE, which is the default.

Details

The default setting is that the time series are normalized with their mean values. Optionally a polynomial of degree 1 or more is first fitted to each time series, whereupon the original time series are normalized by dividing with the fitted values at each point, thus trends in a fold-change sense are assumed. Another option is trend elimination, in which case the same model plus a polynomial: $y = m + a\cos(\omega t) + b\sin(\omega t) + et + ft^2 + \dots$ is fitted to the (possibly normalized) data. In this case, returned p-values still only concern the alternative $c > 0$ as defined above.

Values returned include normalized time series (if normalization is performed), normalization weights (means or polynomial coefficients if polynomial normalization is used), fitted normalized curves, fitted non-normalized curves, a data frame of amplitudes and phases (in radians), p-values according to an F-test (Halberg 1967), Benjamini-Hochberg adjusted p-values, a data frame of approximately 1.96 standard deviations for the amplitude and phase estimates, a matrix of coefficients a and b and possibly c, \dots , the sum square residuals after the fit for each time series, and the covariance matrix for the three independent variables (1 , $\cos(\omega t)$, and $\sin(\omega t)$). The latter can be used in post-processing e.g. to obtain individual p-values for coefficients by t-tests.

Value

A list containing:

means	Vector (if <code>norm.pol=FALSE</code>) or matrix (otherwise) of the means or coefficients of the fitted polynomial used
normts	Matrix of mean-scaled or normalized-by-polynomial time series, same dimensionality as <code>inputts</code>
fit.vals	Matrix of model fitted values to <code>inputts</code>
norm.fit.vals	Matrix of model fitted values to the normalized (trend eliminated or mean scaled) time series
pars	Data frame of estimated amplitudes and phases (in radians, between 0 and 2π)
pvals	Vector of p-values according to an F-test of the model fit against a restricted model (mean-centering only)
qvals	Vector of Benjamini-Hochberg adjusted p-values
ci	Data frame of one-sided approximative 95% (2σ) confidence intervals for the estimated amplitudes and phases
coeffs	Matrix of estimated model parameters a and b
ssr	Vector of sum square residuals for the model fits
df	Scalar if <code>inputts</code> does not contain NAs and Vector otherwise, representing the degrees of freedom of the residuals
ssx	Matrix (3 times 3, if <code>inputts</code> does not contain NAs, a list of such matrices, one for each time series, otherwise 3 times 1)

References

Halberg F, Tong YL, Johnson EA: Circadian System Phase – An Aspect of Temporal Morphology; Procedures and Illustrative Examples. in: The Cellular Aspects of Biorhythms, Springer 1967.

rna.nasc

Menet et al. RNA-Seq Data

Description

Quantification of circadian transcriptional activity in mouse liver.

Format

a data frame with nascent RNA-seq data

Details

The nascent-seq data are thought to reflect transcriptional activities. Raw data was collected from the supplementary material of the original publication cited below. In that study, samples were collected at 6 different times of day in two biological replicates (fields ZT* in the data frame). The field 'mgi_symbol' refers to MGI gene names.

Source

Jerome S Menet, Joseph Rodriguez, Katharine C Abruzzi, Michael Rosbash. Nascent-Seq reveals novel features of mouse circadian transcriptional regulation. eLife 1:e00011 (2012).

Examples

```
data(rna.nasc)
nasc.t <- seq(0, 44, 4)
plot(nasc.t, rna.nasc["Arntl", -1], type="b")
```

rna.polya

Menet et al. RNA-Seq Data

Description

Quantification of circadian transcriptional activity in mouse liver.

Format

a data frame with poly(A)+ RNA-seq data

Details

The poly(A)+-seq data represent mature mRNA abundances. Raw data was collected from the supplementary material of the original publication cited below. In that study, samples were collected at 6 different times of day in two biological replicates (fields ZT* in the data frame). The field 'mgi_symbol' refers to MGI gene names.

Source

Jerome S Menet, Joseph Rodriguez, Katharine C Abruzzi, Michael Rosbash. Nascent-Seq reveals novel features of mouse circadian transcriptional regulation. eLife 1:e00011 (2012).

Examples

```
data(rna.polya)
polya.t <- seq(0, 44, 4)
plot(polya.t, rna.polya["Arntl", -1], type="b")
```

Index

harmonic.regression, [2](#)

rna.nasc, [3](#)

rna.polya, [4](#)