TSrepr R package: Time Series Representations

Peter Laurinec[1]

1 Faculty of Informatics and Information Technologies, Slovak University of Technology in Bratislava

Summary

TSrepr (Laurinec 2018) is R package for time series representations computing. Time series representations are, in other words, methods for dimensionality reduction, feature extraction or preprocessing of time series. They are used for:

- significant reduction of the time series dimensionality
- emphasis on fundamental (essential) shape characteristics
- implicit noise handling
- reducing the dimension will reduce memory requirements and computational complexity of consequent machine learning methods (classification or clustering).

The TSrepr package contains various methods and types of time series representations including Piecewise Aggregate Approximation (PAA), Discrete Fourier Transform (DFT), Perceptually Important Points (PIP), Symbolic Aggregate approXimation (SAX), Piecewise Linear Approximation (PLA) or Clipping. Except for these well-known methods, additional methods suitable for seasonal time series are implemented. These methods are based on the model as for example multiple linear regression, robust regression, generalized additive model or triple exponential smoothing (Laurinec and Lucká 2016). Own feature extraction methods from Clipping representation are also implemented - FeaClip and FeaTrend.

Additional useful functions and methods related to time series representations were also implemented. Package includes functions for normalisations and denormalisations of time series - z-score and min-max methods. It supports simple computation of windowing method, a matrix of representations and forecasting accuracy measures (MAE, RMSE, MAPE, sMAPE etc.).

Methods (functions) were implemented in base R and also in C++ for fast computations. In R, C++ programmes can be written thanks to the package Rcpp (Eddelbuettel and Francois 2011).

The general package for time series representations computations missed until now. The CRAN's time series task view proves previous statement. Packages TSMining (Fan 2015) and jmotif (Senin 2016) both includes implementations of PAA and SAX time series representations methods. However, these packages are mainly focused on motif discovery in time series.

References

Eddelbuettel, Dirk, and Romain Francois. 2011. "Rcpp: Seamless R and C++ Integration." *Journal of Statistical Software, Articles* 40 (8):1–18. https://doi.org/10.18637/jss.v040.i08.

Fan, Cheng. 2015. *TSMining: Mining Univariate and Multivariate Motifs in Time-Series Data*. https://CRAN.R-project.org/package=TSMining.

Laurinec, Peter. 2018. *TSrepr: Time Series Representations*. https://CRAN.R-project.org/package=TSrepr.

Laurinec, Peter, and Mária Lucká. 2016. "Comparison of Representations of Time Series for Clustering Smart Meter Data." In *Lecture Notes in Engineering and Computer Science: Proceedings of the World Congress on Engineering and Computer Science 2016*, 458–63.

Senin, Pavel. 2016. *Jmotif: Time Series Analysis Toolkit Based on Symbolic Aggregate Dicretization, I.e. SAX*. https://CRAN.R-project.org/package=jmotif.