

# Area biplot for time series feature extraction

*Alberto Silva*<sup>1, 2</sup>, *Adelaide Freitas*<sup>1, 2</sup>

<sup>1</sup> University of Aveiro

<sup>2</sup> Center for Research & Development in Mathematics and Applications (CIDMA)

**Abbreviated abstract:** For extracting information about a time series, we have implemented an R package called *areabiplot*, applicable to a Singular Spectrum Analysis trajectory matrix, decomposed through Principal Component Analysis via Partial Least Squares. Our main goal is to visually estimate the periodicity of the time series through groups of quasi-similar triangles employing the so-called SSA area biplot.

## Related publications:

- Gower, J. C. *et al*, Journal of Computational and Graphical Statistics, 19:1, 46-61(2010)
- Silva, A. *et al*, Statistics, Optimization & Information Computing Journal, 8, 346–358 (2020)



CIDMA]

[albertos@ua.pt](mailto:albertos@ua.pt) - 1



3rd Conference on  
Statistics and  
Data Science  
Salvador, Brazil (online)  
October 28-30, 2021

# Background

Singular Spectrum Analysis (SSA):

- It is an exploratory method used in time series (*ts*) analysis.
- Let  $Y$  be a one-dimensional real-valued *ts*:
  - $\mathcal{T}(Y) \rightarrow X_{L \times K}$  (trajectory matrix)
  - $SVD(X) \rightarrow X = \sum_{i=1}^d \sqrt{t_i' t_i} t_i^* p_i' \Leftrightarrow X = T^* \Sigma P'$  (via NIPALS algorithm)
  - Grouping
  - Diagonal averaging

SSA Area Biplot:

- $A = T_2^* \Sigma_2$  and  $B = P_2 \Sigma_2$  (HJ-biplot scheme)
- $a_i'$  is a biplot point ;  $b_i'$  is a biplot arrow
- Pin a row of  $B$  (e.g.,  $b_j'$ ) and rotate all arrows, e.g,  $b_j', j \in \{1, \dots, K\}$
- Connect the endpoints of  $b_j'$  and each  $b_i'$  forming triangles.



# R package *areabiplot*

- ❖ `areabiplot(L =  $\mathbf{T}_2^*$ , S =  $\mathbf{\Sigma}_2$ , R =  $\mathbf{P}_2$ , ord.row, mode = SS, ...)`, where “SS” stands for “similar spread”, the allocation of  $\Sigma$  proposed by Gowen et al. (2010), such that:

$$\mathbf{A} = \left(\frac{L}{K}\right)^{1/4} \mathbf{T}_2^* \mathbf{\Sigma}_2^{1/2} \quad \text{and} \quad \mathbf{B} = \left(\frac{K}{L}\right)^{1/4} \mathbf{P}_2 \mathbf{\Sigma}_2^{1/2}$$

This is the ordinary area biplot (rotation of the points) in which  $\mathbf{a}_i' \mathbf{b}_j = x_{ij}$  is approximated by twice the area of the respective triangle.

- ❖ `areabiplot(L =  $\mathbf{P}_2$ , S =  $\mathbf{\Sigma}_2$ , R =  $\mathbf{P}_2$ , ord.row, mode = HJ, ...)`, in which “HJ” indicates the HJ-biplot scheme to appropriate  $\Sigma$ . This function builds the SSA Area Biplot (arrow rotation), and it uses groups of quasi-similar triangles to estimate the periodicity of a time series. It also provides interpretation for the autocorrelations between L-phased vectors (columns of the trajectory matrix).

The points can be plotted by means of the function `points()`, if it is necessary to the analysis.

# Application

