Notes on SSA

Jonsa Pichat

18 November 2019

1 SSA

Singular spectrum analysis (SSA) can be defined as principal component analysis (PCA) for univariate time series data. As a reminder, PCA consists of the decomposition of a covariance matrix $X^T X$, where $X \in \mathbb{R}^{n \times m}$ is a data matrix (n being the number of samples, m the number of features and the objective is dimension reduction).

On the other hand, SSA consists of the spectral decomposition of a lag-covariance matrix, also $X^T X$, though X (called the trajectory matrix) contains "the complete record of patterns that have occurred within a window of size L" (or embedding dimension). It is defined such that, given a centred time series $y = y_1, \ldots, y_N$ (with $N = L + K - 1$) of finite rank, one has:

$$X = \begin{bmatrix} y_1 & y_2 & y_3 & \cdots & y_L \\ y_2 & y_3 & y_4 & \cdots & y_{L+1} \\ y_3 & y_4 & y_5 & \cdots & y_{L+2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ y_K & y_{K+1} & \cdots & \cdots & y_{L+K} \end{bmatrix}$$

Elser and Tsonis state that "by using lagged copies of a single time series, [we] can define the coordinates of the phase space that will approximate the dynamics of the system from which the time record was sampled".

1.1 Overview

SSA consists of 2 stages:

1. **decomposition**: (i) *embedding* (see trajectory matrix, X above) and (ii) *decomposition*: SVD is applied to X (equivalently, one may do an eigendecomposition of the covariance matrix) to obtain a decomposition into elementary rank-one matrix components.

2. **reconstruction**: (iii) *grouping*: grouped matrix components are created in a clever way and (iv) *diagonal averaging*: those are back-transformed to provide a decomposition of the time series.