Numerical Linear Algebra issues in Singular Spectrum Analysis of time series

Dario Fasino — with E. Bozzo, R. Carniel

University of Udine (Italy)

Genova, 2ggALN'12

Singular Spectrum Analysis (SSA): Introduction

SSA is a quite recent technique for the analysis of experimental time series, based on the SVD of certain Hankel matrices.

Let $x = (x_1, x_2, \ldots, x_\ell)^T$ a finite time series, $\ell = m + n - 1$.
The $m \times n$ Hankel matrix

$$X_{m,n} = \begin{pmatrix} x_1 & x_2 & \cdots & x_n \\ x_2 & x_3 & \cdots & x_{n+1} \\ \vdots & \vdots & \vdots & \vdots \\ x_m & x_{m+1} & \cdots & x_\ell \end{pmatrix}$$

is known as trajectory matrix, $X_{m,n} = \mathcal{T}_{m,n}(x)$.

Singular Spectrum Analysis (SSA): Introduction

SSA is a quite recent technique for the analysis of experimental time series, based on the SVD of certain Hankel matrices.

Let $z \in \mathbb{C}$. Then

$$\mathrm{rank} \begin{pmatrix} 1 & \cdots & z^{n-1} \\ z & \cdots & z^n \\ \vdots & \vdots & \vdots \\ z^{m-1} & \cdots & z^{\ell-2} \end{pmatrix} = 1.$$

Let $x_i = p_k(i)$,
a k-degree algebraic poly. Then

$$\mathrm{rank} \begin{pmatrix} x_1 & \cdots & x_n \\ \vdots & \vdots & \vdots \\ x_m & \cdots & x_\ell \end{pmatrix} = k.$$

Time series made up by trigonometric, algebraic, exponential terms have small rank trajectory matrices.

Singular Spectrum Analysis (SSA): Introduction

SSA idea

Use SVD of trajectory matrices to decompose time series into constant terms, trends, oscillatory components, noise…

Given $x = (x_1, \dots, x_\ell)$ and $\mathcal{I}_1, \dots, \mathcal{I}_k$ a partition of $\{1, \dots, n\}$ do:

1. Build up trajectory matrix: $X = \mathcal{T}_{m,n}(x)$.
2. Compute SVD $X = U\Sigma V^T$; singular triples: (u_i, σ_i, v_i).
3. Group triples: $X^{(k)} = \sum_{i \in \mathcal{I}_k} \sigma_i u_i v_i^T$.
   Note: $\sum_k X^{(k)} = X$.
4. Hankelization (diagonal averaging): $H^{(k)} = \mathcal{H}(X^{(k)})$.
   Note: $\sum_k H^{(k)} = X$.
5. Extract components: $x^{(k)} = \mathcal{T}_{m,n}^{-1}(H^{(k)})$.
   Note: $\sum_k x^{(k)} = x$.

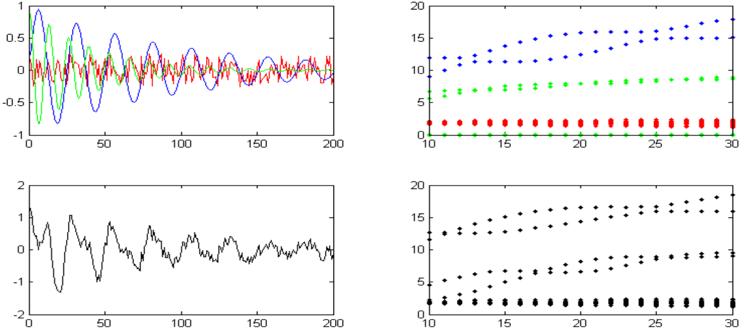

Figure: SSA of a mixture of time series. $\ell = 200$, $n = 10, \ldots, 30$.
Above: individual time series and respective SVs.
Below: composite time series and respective SVs.

Example

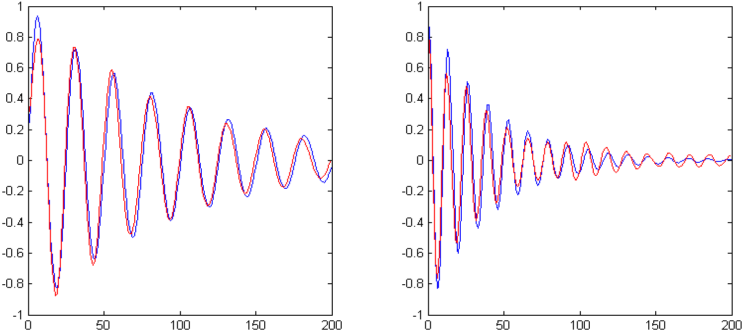

Figure: SSA of a mixture of time series. $\ell = 200$, $n = 30$.
Left: first original component (blue) and its reconstruction (red).
Right: second original component (blue) and its reconstruction (red).

References

N. Golyandina, V. Nekrutkin, A. Zhigljavsky.
Analysis of time series structure. SSA and related techniques.
Chapman & Hall/CRC, 2001.

R. Carniel *et al*.
On the singular values decoupling in the Singular Spectrum Analysis of volcanic tremor at Stromboli.
Nat. Hazards Earth Syst. Sci., 6 (2006), 903–909.

E. Bozzo, R. Carniel, D. F.
Relationship between SSA and Fourier analysis: Theory and application to the monitoring of volcanic activity.
Comp. Math. Appl. 60 (2010), 812–820.

V. Busoni.
Risultati di tipo perturbativo nell'analisi dello spettro singolare di serie temporali.
Tesi di Laurea in Matematica, Università di Udine, 2011.

A motivating problem (and our answer)

April 5, 2003: a major paroxism destroyed a seismic station on the Stromboli volcano.

SSA analysis of the sismogram suggests the presence of a consistent preparatory phase.

What information is conveyed in the SVs of (not so large) trajectory matrices coming from *chaotic* time series?

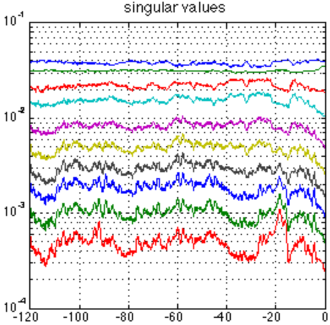

Figure: SSA of a volcanic tremor sismogram.
Singular values of 7200 trajectory matrices $X_{3000,10}$ (smoothed plot; x-axis in hours)

A motivating problem (and our answer)

Our result: The behaviour of SVs mirrors qualitative modifications in the *power spectrum* of the time series.

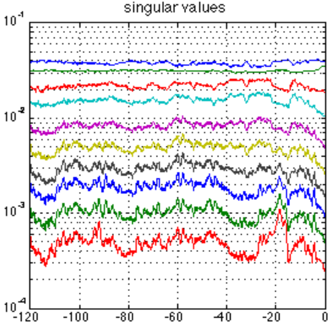

Figure: SSA of a volcanic tremor sismogram.
Singular values of 7200 trajectory matrices $X_{3000,10}$ (smoothed plot; x-axis in hours)

Stationary time series

Definition

The infinite time series $\vec{x} = (x_1, x_2, \dots)$ is called stationary if

$$\forall i, j \geqslant 0 \qquad \lim_{m \to \infty} \frac{1}{m} \sum_{k=1}^{m} x_{i+k} x_{j+k} = R(i-j),$$

where $R : \mathbb{Z} \mapsto \mathbb{R}$ is the *covariance function* of x.

Equivalently, the *covariance matrix*

$$T_n = \lim_{m \to \infty} \frac{1}{m} X_{m,n}^T X_{m,n}$$

exists for all n (and is a Toeplitz matrix).

Stationary time series

Definition

The infinite time series $\vec{x} = (x_1, x_2, \ldots)$ is called stationary if

$$\forall i, j \geqslant 0 \qquad \lim_{m\to\infty} \frac{1}{m} \sum_{k=1}^{m} x_{i+k} x_{j+k} = R(i-j),$$

where $R : \mathbb{Z} \mapsto \mathbb{R}$ is the *covariance function* of x.

By Herglotz theorem, $\vec{x}$ is a stationary time series iff there exists a (unique, bounded) nondecreasing function $\mu(t)$ on $\mathcal{I} = [0, 2\pi]$ such that

$$R(k) = \frac{1}{2\pi} \int_{\mathcal{I}} \mathrm{e}^{\mathrm{i}2\pi kt} \, \mathrm{d}\mu(t), \qquad k \in \mathbb{Z}.$$

Stationary time series

Definition

The infinite time series $\vec{x} = (x_1, x_2, \ldots)$ is called stationary if

$$\forall i, j \geqslant 0 \qquad \lim_{m \to \infty} \frac{1}{m} \sum_{k=1}^{m} x_{i+k} x_{j+k} = R(i-j),$$

where $R : \mathbb{Z} \mapsto \mathbb{R}$ is the *covariance function* of x.

A special case: $\vec{x}$ is called aperiodic or chaotic whenever

$$R(k) = \frac{1}{2\pi} \int_{\mathcal{I}} \mathrm{e}^{\mathrm{i}2\pi kt} f(t)\, \mathrm{d}t, \qquad k \in \mathbb{Z}.$$

The function $f \in L^1(\mathcal{I})$, $f \geqslant 0$, is the spectral density of $\vec{x}$ and the symbol of the Toeplitz matrix sequence $\{T_n\}$.

Asymptotic distributions

Definition

Two triangular sequences $\{\xi_i^{(n)}\}_{i=1\dots n}$ and $\{\zeta_i^{(n)}\}_{i=1\dots n}$, with $n \in \mathbb{N}$, are *equally distributed* (or *asymptotically equidistributed*), $\xi_i^{(n)} \sim \zeta_i^{(n)}$ if, for all continuous functions F having bounded support,

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} \left[F(\xi_i^{(n)}) - F(\zeta_i^{(n)}) \right] = 0.$$

Theorem

Let $\{T_n\}$ be a sequence of Toeplitz matrices whose symbol $f \in L^1(\mathcal{I})$ is Riemann-integrable. Then $\lambda_i(T_n) \sim f(2\pi i/n)$.

Asymptotic distributions

Definition

Two triangular sequences $\{\xi_i^{(n)}\}_{i=1\dots n}$ and $\{\zeta_i^{(n)}\}_{i=1\dots n}$, with $n \in \mathbb{N}$, are *equally distributed* (or *asymptotically equidistributed*), $\xi_i^{(n)} \sim \zeta_i^{(n)}$ if, for all continuous functions F having bounded support,

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} \left[F\big(\xi_i^{(n)}\big) - F\big(\zeta_i^{(n)}\big) \right] = 0.$$

Theorem [Tyrtyshnikov '96]

Let A_n, B_n be $m(n) \times n$ matrices, with $m(n) \geq n$. If

$$\lim_{n\to\infty} \frac{1}{n} \|A_n - B_n\|_F^2 = 0 \qquad \Longrightarrow \qquad \sigma_i(A_n) \sim \sigma_i(B_n).$$

Let $\vec{x} = (x_1, x_2 \ldots)$ be a stationary time series,

$$\widehat{X}_{m,n} = \frac{1}{\sqrt{m}} X_{m,n} \qquad \widetilde{X}_{m,n} = \frac{1}{\sqrt{m}} \begin{pmatrix} x_1 & x_2 & \cdots & x_n \\ x_2 & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & x_1 \\ \vdots & x_m & \cdots & \vdots \\ x_m & x_1 & \cdots & x_{n-1} \end{pmatrix}.$$

Lemma

For any integer sequence $m(n)$ such that $n/m(n) \to 0$ as $n \to \infty$, the singular values of $\widehat{X}_{m(n),n}$ and $\widetilde{X}_{m(n),n}$ are equally distributed.

SSA and Fourier analysis

Indeed:

$$
\begin{aligned}
\frac{1}{n}\|\widehat{X}_{m(n),n} - \widetilde{X}_{m(n),n}\|_F^2 &= \frac{1}{m(n)} \sum_{k=1}^{n} \frac{n-k}{n}(x_k - x_{m(n)+k})^2 \\
&< \frac{2}{m(n)} \sum_{k=1}^{n} x_k^2 + x_{m(n)+k}^2 \\
&= \underbrace{\frac{2n}{m(n)}}_{\to 0} \Bigg(\underbrace{\frac{1}{n} \sum_{k=1}^{n} x_k^2}_{\to R(0)} + \underbrace{\frac{1}{n} \sum_{k=1}^{n} x_{m(n)+k}^2}_{\to R(0)} \Bigg) \to 0.
\end{aligned}
$$

Note: $\vec{x}$ stationary $\Rightarrow$ any trailing subsequence is stationary.

SSA and Fourier analysis

For any fixed window length m, let $x = (x_1, \dots, x_m)^T$,
and let $\hat{x} = (\hat{x}_0, \dots, \hat{x}_{m-1})^T = F_m x$ be its Fourier transform.

Let $\Lambda_m = \mathrm{Diag}(1, \mathrm{e}^{\mathrm{i}2\pi/m}, \dots, \mathrm{e}^{\mathrm{i}2(m-1)\pi/m})$ and $P_m = F_m^H \Lambda_m F_m$.

The matrix $T_{m,n} = \widetilde{X}_{m,n}^T \widetilde{X}_{m,n}$ is Toeplitz, with symbol

$$f_{m,n}(z) = \frac{1}{m} \sum_{k=-n+1}^{n-1} x^T P_m^{-k} x \, \mathrm{e}^{\mathrm{i}kz} = \frac{1}{m} \sum_{k=-n+1}^{n-1} \hat{x}^H \Lambda_m^{-k} \hat{x} \, \mathrm{e}^{\mathrm{i}kz}.$$

Due to the stationarity assumption, $\lim_{m\to\infty} T_{m,n} = T_n$, hence
for “fastly growing” $m(n)$ we have $\lim_{n\to\infty} \frac{1}{n} \| T_{m(n),n} - T_n \|_F^2 = 0$.

$$\sigma_i(\widehat{X}_{m(n),n})^2 \sim \sigma_i(\widetilde{X}_{m(n),n})^2 = \lambda_i(T_{m(n),n}) \sim \lambda_i(T_n) \sim f(2\pi i/n).$$

SSA and Fourier analysis

For any fixed window length m, let $x = (x_1, \ldots, x_m)^T$, and let $\hat{x} = (\hat{x}_0, \ldots, \hat{x}_{m-1})^T = F_m x$ be its Fourier transform.

Let $\Lambda_m = \mathrm{Diag}(1, \mathrm{e}^{\mathrm{i}2\pi/m}, \ldots, \mathrm{e}^{\mathrm{i}2(m-1)\pi/m})$ and $P_m = F_m^H \Lambda_m F_m$.

The matrix $T_{m,n} = \widetilde{X}_{m,n}^T \widetilde{X}_{m,n}$ is Toeplitz, with symbol

$$f_{m,n}(z) = \frac{1}{m} \sum_{k=-n+1}^{n-1} x^T P_m^{-k} x \, \mathrm{e}^{\mathrm{i}kz} = \frac{1}{m} \sum_{k=-n+1}^{n-1} \hat{x}^H \Lambda_m^{-k} \hat{x} \, \mathrm{e}^{\mathrm{i}kz}.$$

Moreover, for any fixed m, n and for $j = 1, \ldots, n$

$$f_{m,n}\left(\frac{2\pi j}{n}\right) = \frac{1}{m} \hat{x}^H \left(\sum_{k=-n+1}^{n-1} \mathrm{e}^{\mathrm{i}k2\pi j/n} \Lambda_m^{-k} \right) \hat{x} = \frac{1}{m} \sum_{i=0}^{m-1} |\hat{x}_i|^2 \eta_{i,j}$$

where $\eta_{i,j} = \sum_{k=-n+1}^{n-1} \mathrm{e}^{\mathrm{i}k2\pi(j/n - i/m)}$.

SSA and Fourier analysis

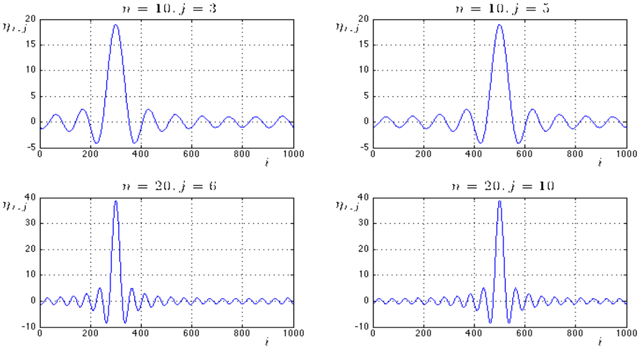

Figure: Plots of the coefficients $\eta_{i,j} = D_n(2\pi(j/n - i/m))$. Here $m = 1000$.

$$D_n(\theta) = \sum_{k=-n+1}^{n-1} \mathrm{e}^{\mathbf{i}k\theta}$$ is the nth *Dirichlet kernel*.

SSA and Fourier analysis

Theorem

If $\vec{x}$ is a stationary time series
(+ assumptions on integrability of f and growth of $m(n)$ as $n \to \infty$)
then the singular values of $\widehat{X}_{m(n),n}$ and $\{f^{1/2}(2\pi j/n)\}_{j=1\dots n}$
are equally distributed.

Each SV of $X_{m,n}$ depends essentially from a portion of
the power spectrum $|\hat{x}_0|^2, \dots, |\hat{x}_{m-1}|^2$ whose width is $\approx m/n$.

Actually, we can replace $f^{1/2}(2\pi j/n)$ with the power bins

$$\varphi_j^{(n)} = \left(\frac{1}{L} \sum_{i=\lfloor (j-1)L \rfloor}^{\lfloor jL \rfloor - 1} |\hat{x}_i|^2 \right)^{1/2} \qquad L = \frac{m(n)}{2n} \qquad j = 1, \dots, n.$$

Numerical example: Pseudorandom time series

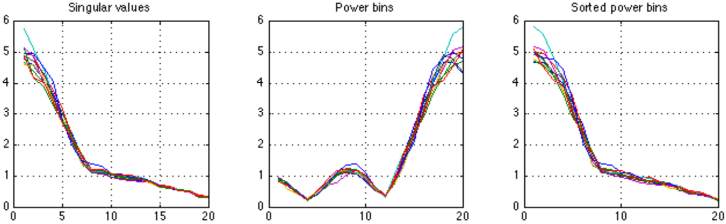

Figure: SSA of 10 pseudorandom time series with spectral density function $f(z) = |1 - 2\cos(z) + 2\cos(2z)|$.

Left: singular values of trajectory matrices $\widehat{X}_{1000,20}$

Center: power bins $\varphi_1^{(20)}, \dots, \varphi_{20}^{(20)}$

Right: power bins $\varphi_i^{(20)}$ in decreasing order.

Numerical example: Stromboli tremor analysis

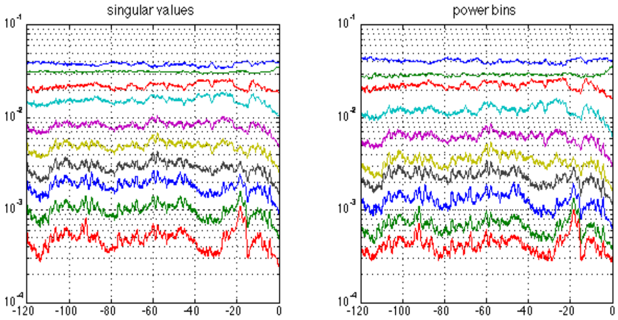

Figure: Analysis of a volcanic tremor signal.

Left: singular values of trajectory matrices $\widehat{X}_{3000,10}$.

Right: power bins $\varphi_i^{(10)}$ in decreasing order.

Separability

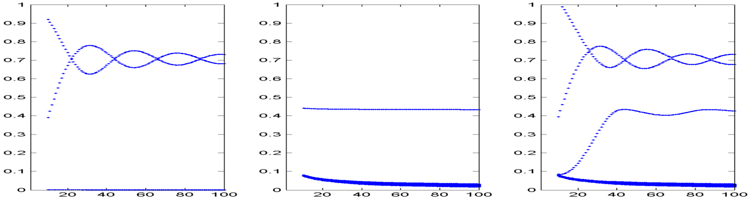

Figure: SSA of two time series and their sum.

In general, the SVs of $\mathcal{T}_{m,n}(x^{(1)})$ and $\mathcal{T}_{m,n}(x^{(2)})$ cannot be recovered from those of $\mathcal{T}_{m,n}(x^{(1)} + x^{(2)})$ — biorthogonality needed.

Problem

To what extent one can recover individual components without biorthogonality?

Separability

Theorem

Let $A, B \in \mathbb{R}^{m\times n}$ *with SVDs* $A = U_A \Sigma_A V_A^T$ *and* $B = U_B \Sigma_B V_B^T$. *Suppose* $\mathrm{rank}(A|B) = \mathrm{rank}(A + B)$. *Moreover, let*

$$\varepsilon_L = \| U_A^T U_B \|, \quad \varepsilon_R = \| V_A^T V_B \|.$$

Then,

$$\frac{1}{\eta} \leqslant \frac{\sigma_i(A+B)}{\sigma_i(A|B)} \leqslant \eta, \qquad \eta = \sqrt{\frac{(1+\varepsilon_L)(1+\varepsilon_R)}{(1-\varepsilon_L)(1-\varepsilon_R)}}.$$

Note: $\varepsilon_L \leqslant \|A^+\| \|B^+\| \|A^T B\|$ and $\varepsilon_R \leqslant \|A^+\| \|B^+\| \|A B^T\|$.

Separability

Theorem

Let $A, B \in \mathbb{R}^{m \times n}$ with SVDs $A = U_A \Sigma_A V_A^T$ and $B = U_B \Sigma_B V_B^T$. Suppose $\mathrm{rank}(A|B) = \mathrm{rank}(A+B)$. Moreover, let

$$\varepsilon_L = \|U_A^T U_B\|, \quad \varepsilon_R = \|V_A^T V_B\|.$$

Then,

$$\frac{1}{\eta} \leqslant \frac{\sigma_i(A+B)}{\sigma_i(A|B)} \leqslant \eta, \qquad \eta = \sqrt{\frac{(1+\varepsilon_L)(1+\varepsilon_R)}{(1-\varepsilon_L)(1-\varepsilon_R)}}.$$

Thank you.