

Therefore,

$$\Delta_a \Phi \leq \mathcal{E} \Phi / a^2. \quad (3.24)$$

The inequality (3.24) shows how the support of the frequency response (with threshold a) is related to the filter power.

3.9.2 SSA Reconstruction as a Linear Filter

Let us return to Basic SSA. Let L be the window length and $(\sqrt{\lambda}, U, V)$ be one of the eigentriples generated by the SVD of the trajectory matrix of $\mathbb{X}_N$ (see Sect. 2.1.1 for notation and definitions). Since the reconstruction operation in Basic SSA is the linear operation, it can be written in matrix form.

Let $K = N - L + 1$, $L^* = \min(L, K)$. Define the diagonal $N \times N$ matrix

$$\mathbf{D} = \text{diag}(1, 2, 3, \dots, L^* - 1, L^*, L^*, \dots, L^*, L^* - 1, \dots, 2, 1)$$

and the $K \times N$ matrix

$$\mathbf{W} = \begin{pmatrix} u_1 & u_2 & u_3 & \cdots & u_L & 0 & \cdots & 0 & 0 & 0 \\ 0 & u_1 & u_2 & u_3 & \cdots & u_L & 0 & \cdots & 0 & 0 \\ \vdots & 0 & \ddots & \ddots & \ddots & \cdots & \ddots & 0 & \cdots & 0 \\ 0 & \cdots & 0 & u_1 & u_2 & u_3 & \cdots & u_L & 0 & \vdots \\ 0 & 0 & \cdots & 0 & u_1 & u_2 & u_3 & \cdots & u_L & 0 \\ 0 & 0 & 0 & \cdots & 0 & u_1 & u_2 & u_3 & \cdots & u_L \end{pmatrix}.$$

Proposition 3.9 *The time series component $\tilde{\mathbb{X}}_N$ reconstructed by the eigentriple $(\sqrt{\lambda}, U, V)$ has the form*

$$\tilde{\mathbb{X}}_N^T = \mathbf{D}^{-1} \mathbf{W}^T \mathbf{W} \mathbb{X}_N^T.$$

Proof First, note that $\mathbf{W} \mathbb{X}_N^T = \mathbf{X}^T U = \sqrt{\lambda} V \in \mathbb{R}^K$. This yields that the vector $\mathbf{W}^T \mathbf{W} \mathbb{X}_N^T$ (of size N) consists of sums along N antidiagonals of $\sqrt{\lambda} U V^T$, the elementary summand of the SVD. Multiplication by $\mathbf{D}^{-1}$ provides the normalization of the sums by the number of summands and therefore by the definition we obtain the elementary reconstructed component. $\square$

Remark 3.11 Let us add the index i to $\tilde{\mathbb{X}}$ to indicate that it corresponds to the i th eigenvector $U = U_i$. Then, evidently, the reconstructed series $\tilde{\mathbb{X}}^{(I)}$ by the set of eigentriples $\{(\sqrt{\lambda_i}, U_i, V_i), i \in I\}$ is equal to the sum of the reconstructed elementary series $\tilde{\mathbb{X}}^{(i)}$. Therefore, the matrix form for $\tilde{\mathbb{X}}^{(I)}$ immediately follows from (3.9).

Proposition 3.9 and Remark 3.11 allow us to describe the reconstruction produced by Basic SSA as an application of a set of linear filters.

Let $L \leq K$. Define the linear filters $\Theta_L, \Theta_{L-1}, \dots, \Theta_1$ and Ψ by their impulse characteristics $\mathbf{h}_{\Theta_L}, \dots, \mathbf{h}_{\Theta_1}$ and $\mathbf{h}_{\Psi}$:

$$\begin{aligned}\mathbf{h}_{\Theta_L} &= (\dots, 0, u_L^\circ, 0, \dots), \\ \mathbf{h}_{\Theta_{L-1}} &= (\dots, 0, u_{L-1}, u_L^\circ, 0, \dots), \\ &\dots \\ \mathbf{h}_{\Theta_1} &= (\dots, 0, u_1, \dots, u_{L-2}, u_{L-1}, u_L^\circ, 0, \dots); \\ \mathbf{h}_{\Psi} &= \text{rev } \mathbf{h}_{\Theta_1} = (\dots, 0, u_L^\circ, u_{L-1}, \dots, u_1, 0, \dots).\end{aligned}$$

Now we can introduce the reconstructing filters $\Phi_k, k = 1, \dots, L$, generated by the vector U :

$$\Phi_k = \Theta_k \circ \Psi / (L - k + 1), \quad (3.25)$$

where ‘ $\circ$ ’ stands for the filter composition, which is equivalent to the convolution ‘ $*$ ’ of the filter impulse characteristics.

Proposition 3.10 *Let $\mathbb{X}_N = (x_1, x_2, \dots, x_N)$. Then the terms of the elementary reconstructed series $\tilde{\mathbb{X}}_N$ corresponding to the eigentriple $(\sqrt{\lambda}, U, V)$ have the following representation:*

- $\tilde{x}_s = (\Phi_{s-K+1}(\mathbb{X}_N))_s$ for $K + 1 \leq s \leq N$;
- $\tilde{x}_s = (\Phi_1(\mathbb{X}_N))_s$ for $L \leq s \leq K$.

This result is the direct consequence of Proposition 3.9. The set of filters providing the reconstruction of $\tilde{x}_s$ for $1 \leq s < L$ can be built in a similar way.

Let us examine two special filters: Φ_1 , which is used for the reconstruction of the middle points of the time series with numbers $L, \dots, K$, and Φ_L , which is used for the reconstruction of the last point only. The former is called *the MPF* (Middle Point Filter) and the latter is referred as *the LPF* (Last Point Filter). In next two sections we consider them separately.

3.9.3 Middle Point Filter

As above, we assume $L \leq K$. According to Proposition 3.10, the MPF Φ_1 acts only at the L th to $(N - L + 1)$ th points. This leads to a limited use of the MPF in the case $L \sim N/2$. The explicit formula for the MPF filter $\Phi_1^{(i)}$ corresponding to the eigenvector $U_i = (u_1, \dots, u_L)^T$ has the following form:

$$\tilde{x}_s = \sum_{j=-(L-1)}^{L-1} \left(\sum_{k=1}^{L-|j|} u_k u_{k+|j|} / L \right) x_{s-j}, \quad L \leq s \leq K. \quad (3.26)$$

It is clearly seen that the order of the MPF is equal to $2L - 1$. Alternative representation of (3.26) is

$$\tilde{x}_s = \sum_{j=1}^L \sum_{l=1}^L u_j u_l x_{s+j-l}/L, \quad L \leq s \leq K. \quad (3.27)$$

Let us enumerate several properties of the MPF $\Phi_1^{(I)}$.

1. The filter $\Phi_1^{(I)}$ is symmetric. Hence the MPF is a zero-phase filter. In particular, the MPF does not change phases of sinusoids.
2. In a particular case of $I = \{i\}$, applying Jensen's inequality, we obtain that the sum of coefficients of the $\Phi_1^{(i)}$ given in (3.27) is not larger than 1:

$$\sum_{j=1}^L \sum_{l=1}^L u_j u_l / L = \left(\sum_{j=1}^L u_j \right)^2 / L \leq \sum_{j=1}^L u_j^2 = 1.$$

3. If the matrix $\mathbf{X}\mathbf{X}^T$ is positive, then the leading eigenvector U_1 is positive too (Perron's theorem) and, therefore, the coefficients of the filter $\Phi_1^{(1)}$ are positive. This is true, for example, in the case of positive time series. If the time series is close to a constant (at the timescale of L), then the coordinates of U_1 will be close one to another and the MPT filter $\Phi_1^{(1)}$ will be close to the so-called triangle (Bartlett) filter. This implies, for instance, that the extraction of trend by Sequential SSA with small L (see Sect. 2.5.5) is similar to the application of a weighted moving average procedure with positive nearly triangular weights.
4. Power of the MPF satisfies the following inequalities.

Proposition 3.11 *Let the filter $\Phi_1^{(i)}$ be the MPF generated by eigenvector $U_i = (u_1, \dots, u_L)^T$. Then $\mathcal{E}\Phi_1^{(i)} \leq 1/L$.*

Proof The proof of the proposition results from the following inequality:

$$\|\mathbf{h}_\Psi * \text{rev } \mathbf{h}_\Psi\| \leq \sum_{j=1}^L |u_j| \cdot \|\mathbf{h}_\Psi\| = \sum_{j=1}^L |u_j| \cdot \|U\| = \sum_{j=1}^L |u_j| \leq \sqrt{L} \|U\| = \sqrt{L}.$$

□

Proposition 3.12 *Let $\Phi_1^{(I)}$ be the MPF generated by eigenvectors $\{U_i, i \in I\}$ where $|I| = r$. Then $\mathcal{E}\Phi_1^{(I)} \leq r^2/L$.*

Proof By linearity of the grouping operation, $\Phi_1^{(I)} = \sum_{i \in I} \Phi_1^{(i)}$, and therefore, by Proposition 3.11 we have:

$$\mathcal{E}\Phi_1^{(I)} = \left\| \sum_{i \in I} \mathbf{h}_{\Phi_1^{(i)}} \right\|^2 \leq \left(\sum_{i \in I} \|\mathbf{h}_{\Phi_1^{(i)}}\| \right)^2 = \left(\sum_{i \in I} \sqrt{\mathcal{E}\Phi_1^{(i)}} \right)^2 \leq r^2/L.$$

□

5. A direct consequence of Proposition 3.12 and inequality (3.24) is the inequality $\Delta_a \Phi_1^{(I)} \leq r^2/(a^2 L)$. This means that for any threshold a , the support of filter frequency response tends to 0 as $L \rightarrow \infty$. This effect is clearly seen in Fig. 2.22 (Sect. 2.4.3) showing the smoothing effect of Basic SSA.
6. Let us define for $\omega \in (-0.5, 0.5]$

$$g_U(\omega) = \frac{1}{L} \left| \sum_{j=1}^L u_j e^{-i2\pi\omega j} \right|^2. \quad (3.28)$$

The function g_U is closely related to the periodogram Π_u^L introduced in (2.10) of Sect. 2.3.1.1: $g_U(k/L) = L \Pi_u^L(k/L)/2$ for $0 < k < N/2$ and $g_U(k/L) = L \Pi_u^L(k/L)$ otherwise. It appears that the frequency response of the MPF is almost the same as the periodogram of the vector U .

Proposition 3.13 *Let A_{Φ_1} be the frequency response of the MPF filter Φ_1 . Then $g_U(\omega) = A_{\Phi_1}(\omega)$.*

Proof Recall that $\Phi_1 = \Theta_1 \circ \Psi / L$ where $\mathbf{h}_\Psi = (\dots, 0, u_L^\circ, u_{L-1}, \dots, u_1, 0, \dots)$ and $\mathbf{h}_{\Theta_1} = \text{rev } \mathbf{h}_\Psi$. Also, from the theory of linear filters [20] we have $A_{\Phi_1 \circ \Psi}(\omega) \equiv A_{\Phi_1}(\omega) A_\Psi(\omega)$. Then

$$A_{\Phi_1}(\omega) = \frac{1}{L} \left| \sum_{j=0}^{L-1} u_{L-j} e^{-i2\pi\omega j} \right| \cdot \left| \sum_{j=0}^{1-L} u_{L+j} e^{-i2\pi\omega j} \right| = \frac{1}{L} \left| \sum_{j=1}^L u_j e^{-i2\pi\omega j} \right|^2. \quad \square$$

7. It follows from Proposition 3.13 that for SSA identification and interpretation of the SVD components, the periodogram analysis of eigenvectors can be very helpful. Also, an automatic identification of components introduced in Sect. 2.4.5 is based on properties of periodograms of eigenvectors and therefore can also be expressed in terms of the frequency response of the MPF.

3.9.4 Last Point Filter and Forecasting

The last-point filter (LPF) is not really a filter as it is used only for the reconstruction of the last point: $\tilde{x}_N = \sum_{i=0}^{L-1} u_L u_{i+1} x_{N-i}$. The reconstruction by the eigentriples with numbers from the set I has the following form:

$$\tilde{x}_N^{(I)} = \sum_{k=0}^{L-1} \left(\sum_{i \in I} u_L^{(i)} u_{k+1}^{(i)} \right) x_{N-k}. \quad (3.29)$$

However, it is the only reconstruction filter that is causal. This has two consequences. First, the LPF of a finite-rank series is closely related to the LRR governing and forecasting this time series. Second, the so-called Causal SSA (or last-point SSA) can be constructed by means of the use of the last reconstructed points of the accumulated data. Since in the Causal SSA the LPF is applied many times, the study of properties of LPF is important.

Let the signal $\mathbb{S}_N$ has rank r and is governed by an LRR. Unbiased causal filters of order L and linear recurrence relations of order $L-1$ are closely related. In particular, if the causal filter is given by $s_j = \sum_{k=0}^{L-1} a_{L-k} s_{j-k}$ and $a_L \neq 1$, then this filter generates the LRR of order $L-1$: $s_j = \sum_{k=1}^{L-1} c_{L-k} s_{j-k}$, where $c_k = a_k/(1-a_L)$.

Similar to the minimum-norm LRR, the minimum-power filters can be considered. It appears that the LPF has minimal power among all unbiased filters. This follows from the relation between LRRs and causal filters. Denote by P_r the orthogonal projector on the signal subspace. The LPF has the form $s_N = (P_r S_K)_L = A^T S_K$, where $A = P_r \mathbf{e}_L$, while the min-norm LRR is produced by the last-point filter, i.e. $R = A/(1-a_L)$.

In the general case, the filter (3.29) can be rewritten as $\tilde{x}_N = (P_r X_K)_L$, where X_K is the last L -lagged vector, P_r is the projector on SSA estimate of the signal subspace $\text{span}(U_i, i \in I)$. Formally applying this filter to the whole time series, we obtain the series of length K consisting of the last points of the reconstructed lagged vectors $\tilde{X}_k$.

Note that if we use other estimates of the signal subspace, then we obtain other versions of the last-point filter.

3.9.5 Causal SSA (Last-Point SSA)

Let $\mathbb{X}_\infty = (x_1, x_2, \dots)$ be an infinite series, $\mathbb{X}_M = (x_1, \dots, x_M)$ be its subseries of length M , L be fixed, $\mathcal{L}(M)$ be a subspace of $\mathbb{R}^L$, $P(M)$ be a projector to $\mathcal{L}(M)$, $A(M) = P_r(M) \mathbf{e}_L$, $K = K(M) = M - L + 1$.

Introduce the series $\check{\mathbb{X}}_\infty$ as follow. Define $(\check{\mathbb{X}}_\infty)_M = (P(M) X_{K(M)})_L = A(M)^T X_{K(M)}$, where $X_{K(M)}$ is the last L -lagged vector of $\mathbb{X}_M$. Thus, $(\check{\mathbb{X}}_\infty)_M$ is a linear combination of the last L terms of $\mathbb{X}_M$ with coefficients depending on M ; that is, $\check{\mathbb{X}}_\infty$ can be considered as a result of application of a sequence of different causal filters to $\mathbb{X}_\infty$.

If $\mathcal{L}_r(M) = \mathcal{L}(M) = \text{span}(U_1(M), \dots, U_r(M))$, where $U_1(M), \dots, U_r(M)$ are the signal eigenvectors produced by SSA with window length L applied to $\mathbb{X}_M$, then this sequence of causal filters is called Causal SSA. In this case, $(\check{\mathbb{X}})_M$ is equal to the last point of SSA reconstruction $\tilde{\mathbb{X}}_M$, which in turn is equal to the last coordinate of the projection of the last lagged vector of $\mathbb{X}_M$ to $\mathcal{L}_r(M)$.

Given that $\mathcal{L}_r(M)$ is used as an estimate of the signal subspace, M should be large enough to provide a good estimate. Therefore, we need to introduce a starting point M_0 (with $M_0 > L$) in the Causal SSA and consider $M \geq M_0$ only.

Since the result of application of the Causal SSA is a sequence $\check{\mathbb{X}}_\infty$ built from SSA reconstructions of the last points of the subseries, the Causal SSA can be called Last-point SSA.

Remark 3.12 Let us fix N, L and consider $\mathbb{X}_N$, the corresponding $U_1, \dots, U_r$ and $\mathcal{L}(M) = \text{span}(U_1, \dots, U_r)$ for any M . Then the series $\check{\mathbb{X}}_N$ is the result of application of the last-point filter to $\mathbb{X}_N$. Assuming that the estimates of the signal subspace on the base of $\mathbb{X}_M$ are stable for large enough M , we can conclude that the result of the Causal SSA will be close to the result of the last-point filter (LPF).

Note also that in the considered particular case, $\check{\mathbb{X}}_N$ (more precisely, its last K points; the first $L - 1$ points are not defined or can be set to zero) coincides with the last row of the reconstructed matrix $\hat{\mathbb{X}}$ of the series $\mathbb{X}_N$. That is, $\check{\mathbb{X}}_N$ is similar to the result of reconstruction before the diagonal averaging is made.

Causality yields the following relation: under the transition from $\mathbb{X}_M$ to $\mathbb{X}_{M+1}$, the first M points of the $\check{\mathbb{X}}_{M+1}$ coincide with $\check{\mathbb{X}}_M$. This is generally not true if we consider the reconstructions $\check{\mathbb{X}}_M$ for $\mathbb{X}_M$ obtained by the conventional SSA. The effect $(\check{\mathbb{X}}_M)_j \neq (\check{\mathbb{X}}_{M+1})_j, j \leq M$, is called ‘redrawing’. For real-life time series we usually have redrawing for all j and the amount of redrawing depends on j . Redrawing of only a few last points is usually of interest. Moreover, redrawing of local extremes of the series is more practically important than redrawing of regular series points. Small values of redrawing indicate stability of SSA decompositions and hence stability of time series structure. The amount of redrawing can be assessed by visual examination or measured using the variance of the redrawings at each time moment of interest. These variances can be averaged if needed.

Generally, the reconstruction has no delay (at least, the middle-point filter has zero phase shift). On the other hand, delays in the Causal SSA are very likely. In a sense, a redrawing in SSA is transferred to a delay in the Causal SSA. If $\mathcal{L}_r$ corresponds to the exactly separated component of the time series $\mathbb{X}$, then SSA has no redrawing and the Causal SSA has no delay. In conditions of good approximate separability the redrawing is almost absent and the delay is small.

Example

Let us demonstrate the described effects on the ‘S&P500’ example introduced in Sect. 2.5.1. Figure 3.7 shows the result of the Causal SSA with window length $L = 30$, $\mathcal{L}_r(M) = \text{span}(U_1(M), U_2(M))$ and $M_0 = 200$. The delay is clearly seen. If we consider non-causal (Basic) SSA reconstructions of cumulative subseries, then the redrawing takes place (Fig. 3.8). This redrawing increases in the points of local maximums and minimums.

Finally, let us note that if the direction of change of the time series is of primary concern (rather than the values themselves), then, instead of taking differences of the Causal SSA series, it may be worthwhile considering the time series which consists of differences of the last two points of reconstructions. The result should be expected

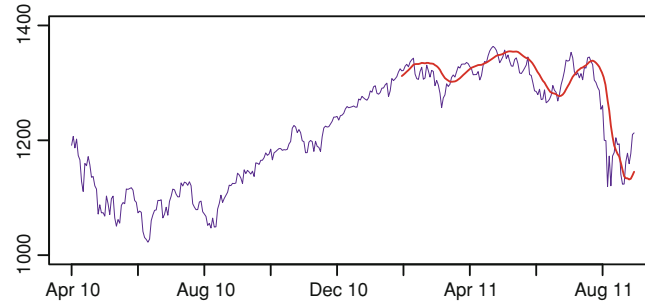


Fig. 3.7 S&P500: causal SSA

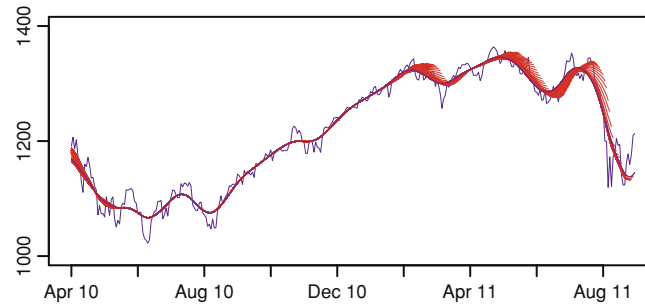


Fig. 3.8 S&P500: non-causal SSA with redrawing

to be more stable since the reconstruction of the next to last point has better accuracy than that of the last point.

References

1. Badeau R, David B, Richard G (2004) Selecting the modeling order for the ESPRIT high resolution method: an alternative approach. In: Proceedings of the IEEE ICASSP, vol 2, pp 1025–1028
2. Badeau R, Richard G, David B (2008) Performance of ESPRIT for estimating mixtures of complex exponentials modulated by polynomials. *IEEE Trans Signal Process* 56(2):492–504
3. Barkhuijsen H, de Beer R, van Ormondt D (1987) Improved algorithm for noniterative time-domain model fitting to exponentially damped magnetic resonance signals. *J Magn Reson* 73:553–557
4. Beckers J, Rixen M (2003) EOF calculations and data filling from incomplete oceanographic data sets. *Atmos Ocean Technol* 20:1839–1856
5. Bozzo E, Carniel R, Fasino D (2010) Relationship between singular spectrum analysis and Fourier analysis: theory and application to the monitoring of volcanic activity. *Comput Math Appl* 60(3):812–820
6. Cadzow JA (1988) Signal enhancement: a composite property mapping algorithm. *IEEE Trans Acoust* 36(1):49–62