BOUNDARY CONDITIONS FOR SYMMETRIC BANDED TOEPLITZ MATRICES: AN APPLICATION TO TIME SERIES ANALYSIS

Alessandra Luati
Dip. Scienze Statistiche
University of Bologna

Tommaso Proietti
S.E.F. e ME. Q.
University of Rome "Tor Vergata"

⋆ ⋆ ⋆ ⋆ ⋆

DUE GIORNI DI ALGEBRA LINEARE NUMERICA, BOLOGNA, 6-7 MARZO 2008

This talk concerns the **spectral properties of matrices associated with linear filters** for the estimation of the underlying trend of a time series.

These matrices are **finite approximations of infinite symmetric banded Toeplitz (SBT) operators** subject to boundary conditions.

The interest lies in the fact that the **eigenvectors can be interpreted as the latent components** of a time series that the filter smooths through the eigenvalues.

In this study, analytical results on the eigenvalues and eigenvectors of matrices associated with trend filters are derived by interpreting the latter as perturbations of matrices belonging to algebras with known spectral properties, such as **the circulant and the reflecting** τ_{11}.

The results allow to design new estimators based on **cut off eigenvalues**, which are less variable and almost equally biased as the original estimators.

Signal extraction of a time series

Time series additive models

$$y_t = \mu_t + \epsilon_t, t = 1, \ldots, n$$

y_t observed time series

μ_t trend component or signal, smooth function of time

ϵ_t irregulars or noise, zero mean stationary stochastic process.

The aim is to estimate μ_t using the available observations.

Smoothing methods like local polynomial regression may serve to this purpose.

Local polynomial regression methods

The basic assumption is that μ_t locally approximated by a p-degree polynomial function of the time distance j between y_t and neighboring y_{t+j}

$$\mu_{t+j} \approx m_{t+j}$$

with

$$m_{t+j} = \beta_0 + \beta_1 j + \cdots + \beta_p j^p, j = 0, \pm 1, \ldots, \pm h.$$

The parameters $\beta_0, ..., \beta_p$ are usually estimated by ordinary or weighted least squares. Our interest lies on $\hat{\beta}_0 = \hat{m}_t$.

Once fixed (i) degree of fitting polynomial, (ii) shape of kernel weighting function, (iii) bandwidth, all the estimators become linear combinations of the input data and are called filters or smoothers.

Weighted least squares estimation

Provided that $\mathbf{y}_{2h+1} = [y_{t-h}, \cdots, y_t, \cdots, y_{t+h}]'$, $2h \geq d$,

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{K}\mathbf{X})^{-1}\mathbf{X}'\mathbf{K}\mathbf{y}_{2h+1}$$

where $\mathbf{X}$ design matrix and $\mathbf{K}$ diagonal weighting matrix.

The trend estimate at time t is

$$\hat{m}_t = \hat{\beta}_0 = \mathbf{e}_1'\hat{\boldsymbol{\beta}} = \mathbf{e}_1'(\mathbf{X}'\mathbf{K}\mathbf{X})^{-1}\mathbf{X}'\mathbf{K}\mathbf{y} = \mathbf{w}'\mathbf{y} = \sum_{j=-h}^{h} \mathsf{w}_j y_{t-j},$$

where $\mathbf{e}_1' = [1, 0, \ldots, 0]$

Specifically:

- **In the interior**, i.e. for $t = h+1, ..., n-h+1$, the trend estimate at time t is

$$\hat{m}_t = \sum_{j=-h}^{h} \mathsf{w}_j y_{t-j}.$$

- **At the boundaries**, asymmetric filters are obtained. For the end of the sample, $t = n-h, ..., n$, crucial in current analysis,

$$\hat{m}_t = \sum_{j=-q}^{h} \mathsf{w}_j y_{t-j},$$

where $q = 0, 1, ...h-1$.

Though the resulting trend estimate is local since it depends only on the subset of the observations that belong to the neighborhood of time t, it can be represented for all $t = 1, ..., n$ as

$$\hat{m}_t = \sum_{j=1}^{n} \mathsf{w}_{tj} y_j .$$

In matrix form, any linear smoother can be represented by a square matrix

$$\begin{aligned} \mathbf{S} \quad &: \quad \mathbf{R}^n \to \mathbf{R}^n \\ \mathbf{y} \quad &\longmapsto \quad \widehat{\mathbf{m}} = \mathbf{S}\mathbf{y}. \end{aligned}$$

$\mathbf{S}$ can be interpreted as a **finite approximation of a symmetric banded Toeplitz operator (SBT)** subject to boundary conditions. In general, $\mathbf{S}$ is centrosymmetric but not symmetric.

The analysis of the properties of $\mathbf{S}$ provides useful information on the local polynomial regression estimator applied to the data to get the smoothed trend.

Spectral analysis

For $n \to \infty, \mathbf{S} \to \mathbf{T}_\infty$, SBT, with real eigenvalues and eigenvectors. In order to understand, *let us suppose* that the spectrum of $\mathbf{S} \to \mathbf{T}_\infty$ is discrete so that

$$\lambda_1 \geq \lambda_2 \geq ... \geq \lambda_n \geq ...$$

then the eigenvectors $\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_n, ...,$ are time series that the filter expands, $\lambda_i > 1$, leaves unchanged, $\lambda_i = 1$, shrinks, $\lambda_i < 1$, or suppresses, $\lambda_i = 0$.

The eigenvectors can be viewed as latent component of any time series that the filter smooths through the eigenvalues.

In fact, let

$$\mathbf{y} = \alpha_1\mathbf{x}_1 + \alpha_2\mathbf{x}_2 + ... + \alpha_n\mathbf{x}_n + ...$$

then

$$\mathbf{Sy} = \sum_{i=1}^{k} \lambda_i\alpha_i\mathbf{x}_i + \sum_{i=k+1}^{\infty} \lambda_i\alpha_i\mathbf{x}_i.$$

Infinite dimension

The non null elements of the SBT operator $\mathbf{T}_\infty$ are the Fourier coefficients of the trigonometric polynomial (the symbol of the matrix, Grenander and Szegö, 1958)

$$H(\nu) = \sum_{d=-h}^{h} \mathsf{w}_d e^{\imath \nu d}$$

and

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} \lambda_i = \frac{1}{2\pi} \int_0^{2\pi} H(\nu) d\nu$$

with

$$\lambda_1 \leq \max H(\nu), \lambda_n \geq \min H(\nu).$$

$H(\nu)$ is the transfer function of the filter evaluated at the frequency ν, radians.

Finite dimension

In finite dimension, the analytical form of eigenvalues and eigenvectors is known only for **few classes of matrices**, that are the tridiagonal SBT and matrices belonging to the Circulant, the Hartley and the generalised Tau algebras.

Interpreting a smoothing matrix as the sum of a matrix belonging to one of these algebras **plus a perturbation occurring at the boundaries**, approximate results on eigenvalues and eigenvectors of $\mathbf{S}$ can be obtained.

The size of the perturbation depends on the (a) **matrix algebra** and on the (b) **boundary conditions**.

We have considered (a) the circulant and the **reflecting τ_{11} algebras** as well as (b) **asymmetric filters** that approximate a given two sided symmetric (Henderson) filter according to a minimum mean square revision error criterion subject to constraints.

Reflecting boundary conditions

The hypothesis is that the first missing observation is replaced by the last available observation, the second missing observation is replaced by the previous to the last observation and so on, that for a two-sided $2h+1$-term filter corresponds to the real time filter, last row of a matrix $\mathbf{H} \in \tau_{11}$ **cosines algebra**,

$$\{0, 0, ..., 0, \mathsf{w}_h, \mathsf{w}_{h-1} + \mathsf{w}_h, ..., \mathsf{w}_1 + \mathsf{w}_2, \mathsf{w}_0 + \mathsf{w}_1\}.$$

Additive decomposition

$$\mathbf{S} = \mathbf{H} + \mathbf{\Delta_H} = \begin{bmatrix} \mathbf{H}^a_{(h\times 2h)} + \mathbf{\Delta_H} & & \mathbf{O}_{(h\times n-2h)} \\ & \mathbf{H}^s_{(n-2h\times n)} & \\ \mathbf{O}_{(h\times n-2h)} & & \mathbf{H}^{a*}_{(h\times 2h)} + \mathbf{\Delta_H} \end{bmatrix}$$

On the matrix $\mathbf{H}$ (Bini and Capovani, 1983, Proposition 2.2).

$$\mathbf{H} = \sum_{j=1}^{n} c_j \mathbf{T}_{\psi\varphi}^{j-1}$$

where

$$\mathbf{T}_{\psi\varphi} = \begin{bmatrix} \psi & 1 & 0 & \cdots & 0 \\ 1 & 0 & 1 & \ddots & 0 \\ 0 & 1 & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & 0 & 1 \\ 0 & \ldots & 0 & 1 & \varphi \end{bmatrix}$$

with $\psi, \varphi = 0, \mathbf{1}, -1$, and $\mathbf{c}$ is the solution of the upper triangular system $\mathbf{Qc} = \mathbf{h}$ where $\mathbf{h}'$ is the first row of $\mathbf{H}$ and $\mathbf{Q}$ is the matrix whose j-th column equals the first column of $\mathbf{T}_{\psi\varphi}^{j-1}$.

Eigenvalues

Theorem *Let* $\mathbf{S}$ *be an* $n \times n$ *smoothing matrix associated with the symmetric filter* $\{\mathsf{w}_{-h}, ..., \mathsf{w}_0, ..., \mathsf{w}_h\}$, *and let* $\mathbf{H}$ *be the corresponding matrix in* τ_{11}. *Hence,* $\forall \lambda \in \sigma(\mathbf{S})$, $\exists i \in \{1, 2, .., n\}$ *such that*

$$|\lambda - \xi_i| \leq \delta_H$$

where

$$\xi_i = \sum_{j=1}^{h+1} \left(2 \cos \frac{(i-1)\pi}{n}\right)^{j-1} \left[\mathsf{w}_{j-1} + \sum_{q=0}^{\left\lfloor \frac{h-j-1}{2} \right\rfloor} \frac{(-1)^{q+1}(j)_q}{(q+1)!}(j+2q+1)\mathsf{w}_{j+2q+1}\right]$$

and $\delta_H = \|\mathbf{S} - \mathbf{H}\|_2$.

Note that $\|\mathbf{A}\|_2 = \sqrt{\rho(\mathbf{A}'\mathbf{A})}$ and $\rho(\mathbf{A})$ is the spectral radius of $\mathbf{A}$.

Figure 1: Transfer function of the **symmetric Henderson filter**, $h = 6$, $\nu \in [0, \pi]$ (line) and eigenvalues of the associated reflecting matrix $\mathbf{H}$ (crosses), $n = 51$.

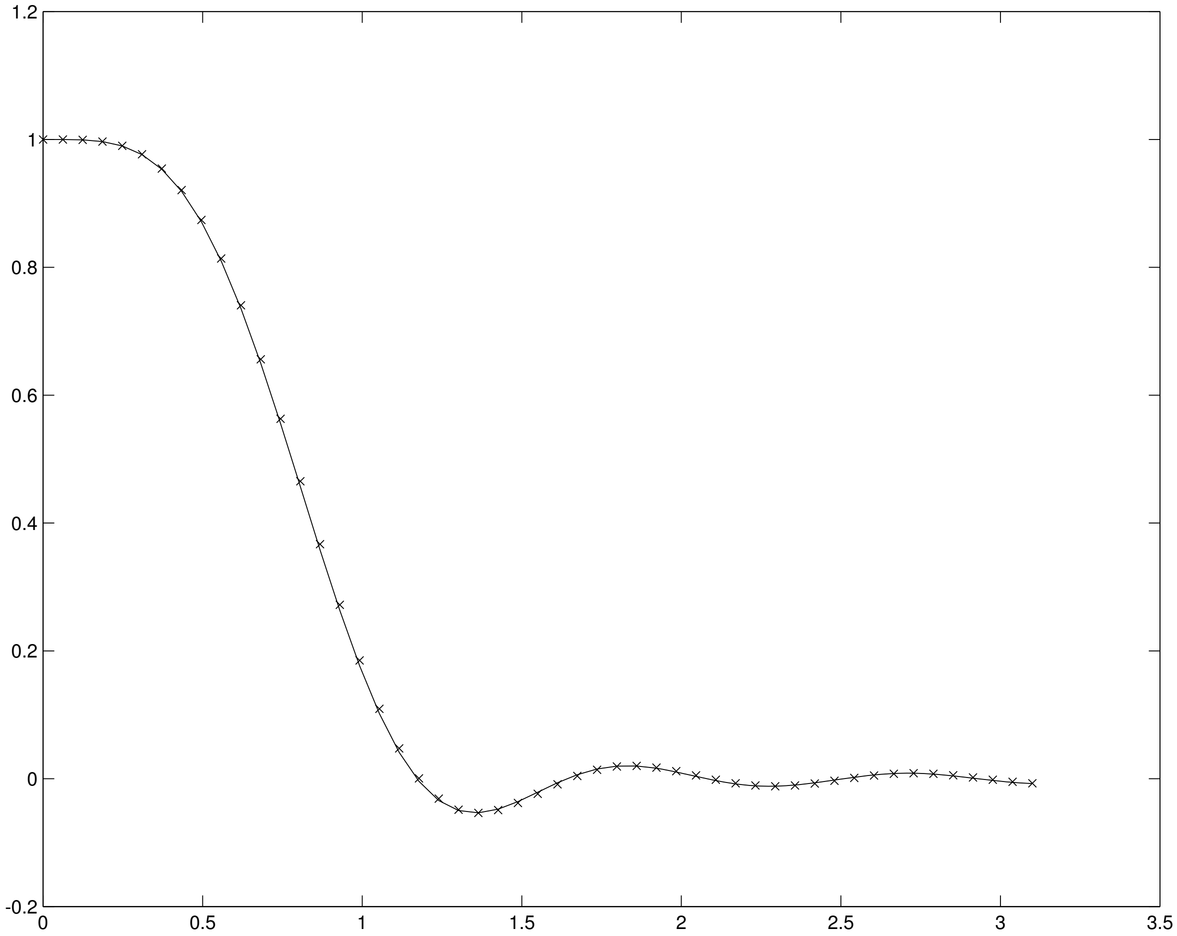
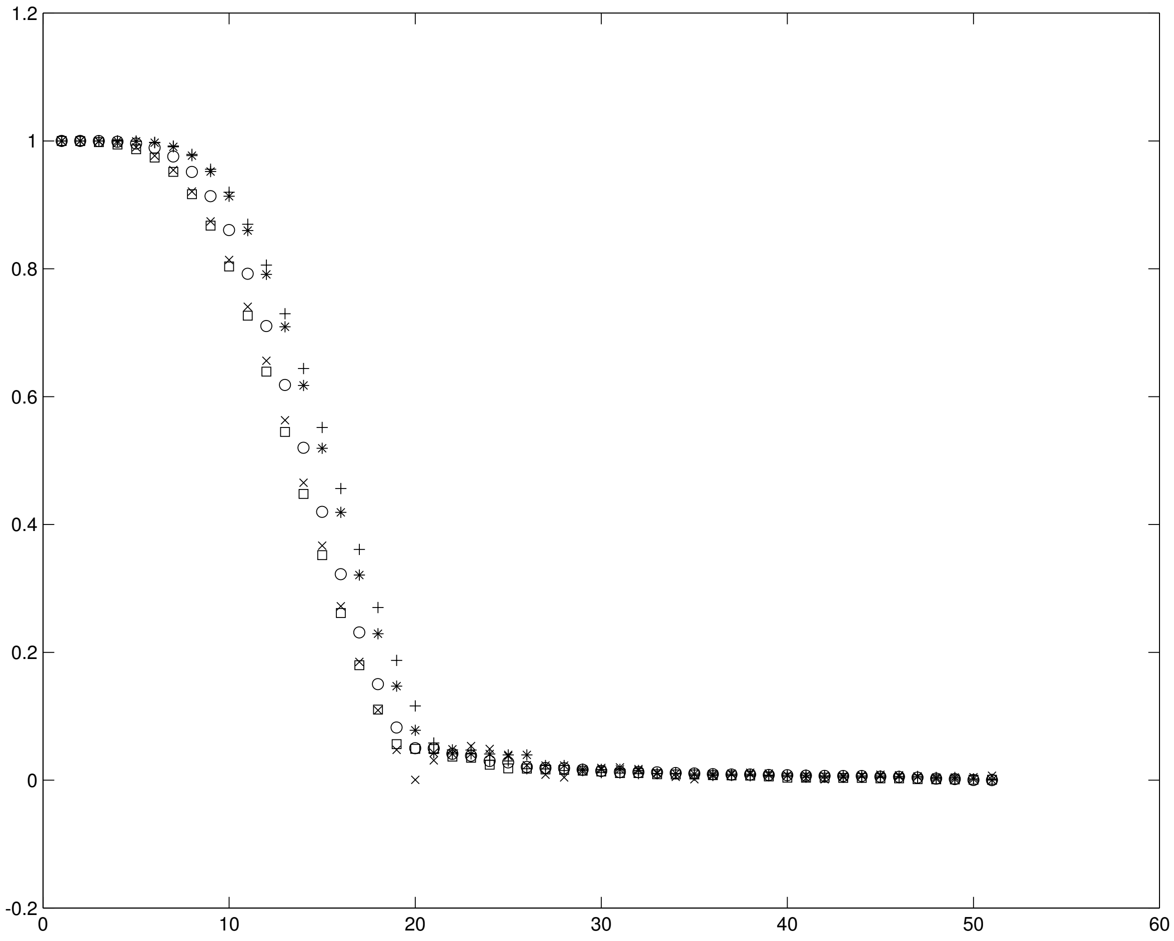


Figure 2: Absolute eigenvalue distributions of $\mathbf{H}$ (crosses) with **asymmetric** Musgrave-LC (squares), QL (circles), CQ (stars), DAF (pluses) filters.



On the size of the perturbation

The quantity δ_H provides an **upper bound** to the size of the perturbation and measures how much the (absolute) eigenvalue distribution of any $\mathbf{S}$ moves away from that of the corresponding reflecting (gain function).

The size of the perturbation **does not depend on** n, in that the $n - 2h$ central rows of the matrix $\mathbf{\Delta_H}$ are all null.

On the other hand, it is **highly influenced by the real time filter** (last row of $\mathbf{S}$), applied to estimate the trend at time t using the available observations up to and including t.

Another factor that highly affects the size of the perturbation **and the overall variance of the trend estimates** is the **algebraic multiplicity of the eigenvalue** $\lambda = 1$, that we have shown to be equal to the degree of the polynomial that the filter is capable to reproduce.

On reflecting versus circulant algebras

All the operators belonging to a τ algebras have real eigenvalues and eigenvectors. All the computations related to this class can be therefore done in real arithmetic.

Circulant-to-Toeplitz corrections produce perturbations that are not smaller than Tau-to-Toeplitz corrections.

The reflecting matrices have n distinct eigenvalues contra the at most $\frac{n-1}{2} + 1$ of the circulant.

On reflecting versus circular boundary conditions

The reflecting hypothesis is more appropriate than that of a circular process when the signal is a non stationary function of time.

Eigenvectors

In general, the analytical expression of the eigenvectors of a smoothing matrix cannot be derived using the perturbation theory. Exact result: there exist up to $p+1$ polynomial eigenvectors. Otherwise,

$$\mathbf{y} = \theta_1 \mathbf{z}_1 + \theta_2 \mathbf{z}_2 + ... + \theta_n \mathbf{z}_n$$

where $\boldsymbol{\theta} = [\theta_1, ..., \theta_n]'$ and $\mathbf{z}_i$ are the real and orthogonal eigenvectors of $\mathbf{H}$,

$$\mathbf{z}_i = k_i \left[\cos \frac{(2j-1)(i-1)\pi}{2n} \right]_j, j = 1, 2, ..., n$$

associated with the eigenvalues ξ_i. Then

$$\mathbf{Sy} = \sum_{i=1}^{n} \theta_i \xi_i \mathbf{z}_i + \sum_{i=1}^{n} \theta_i \boldsymbol{\Delta}_{\mathbf{H}} \mathbf{z}_i$$

where $\boldsymbol{\Delta}_{\mathbf{H}} \mathbf{z}_i$ is a null vector except but for the first and last h coordinates.

Filter design

Reverting to matrix notation, decomposing $\mathbf{S} = \mathbf{H} + \mathbf{\Delta_H}$ and $\mathbf{H} = \mathbf{Z}\mathcal{X}\mathbf{Z}'$, where $\mathcal{X} =$ diag$\{\xi_1, \xi_2, ..., \xi_n\}$, and writing $\mathbf{y} = \mathbf{Z}\boldsymbol{\theta}$,

$$\begin{aligned} \mathbf{S}\mathbf{y} &= \mathbf{Z}\mathcal{X}\boldsymbol{\theta} + \mathbf{\Delta_H}\mathbf{Z}\boldsymbol{\theta} \\ &\approx \mathbf{Z}\mathcal{X}_k\boldsymbol{\theta} + \mathbf{\Delta_H}\mathbf{Z}\boldsymbol{\theta} \end{aligned}$$

where $\mathcal{X}_k$ is the diagonal matrix obtained by **replacing by zeros the eigenvalues of $\mathbf{H}$ that are smaller than a cut-off eigenvalue** ξ_k.

Turning to the original coordinate system and **arranging the boundaries**, we get the new estimator

$$\mathbf{S}_k = \mathbf{H}_k + \mathbf{\Delta_H}$$

where $\mathbf{H}_k$ is the matrix with boundaries equal to those of $\mathbf{H}$ and interior equal to that of $\mathbf{Z}\mathcal{X}_k\mathbf{Z}'$.

In practice:

- given a symmetric filter $\{\mathrm{w}_{-h}, \mathrm{w}_{-h+1}, \ldots, \mathrm{w}_0, \ldots, \mathrm{w}_h\}$
- construct $\mathbf{H} \in \tau_{11}$, reflecting
- obtain the spectral decomposition of $\mathbf{H} \rightarrow$ **possible because of analytically known real eigenvalues orthogonal eigenvectors**
- replace $\mathbf{H} = \mathbf{Z}\mathcal{X}\mathbf{Z}'$ by $\mathbf{H}_{(k)} = \mathbf{Z}\mathcal{X}_k\mathbf{Z}'$
- adjust the boundaries of $\mathbf{H}_{(k)}$ with suitable asymmetric filters to obtain $\mathbf{S}_k$.

Properties

It is proved that the new estimator $\mathbf{S}_k$ so obtained has **smaller variance** in the interior and **almost** equal bias than the original $\mathbf{S}$.

Choice of k

Further balancing of the trade-off between bias and variance of the filter. The trend in the interior is made smoother without sensibly increasing the bias. Alternatives:

Minimum distance of the **eigenvalue distribution** of $\mathbf{H}$ with that of the ideal low pass filter having first k eigenvalues equal to one and and last $n-k$ equal to zero.

Minimum distance between the **transfer functions** of the symmetric filter and of the ideal low-pass filter. If the cut-off frequency is $\nu = \nu_k$, then the cut-off time is

$$k = \frac{\nu_k n}{\pi}.$$

Graphical analysis. Having plotted the eigenvalue distribution, a suitable cut-off eigenvalue may be directly viewed.

Illustrations

We have chosen k such that

$$\|\mathbf{i}_{(k)} - \boldsymbol{\xi}\| \text{ is minimum}$$

where $\mathbf{i}_{(k)} = [1, 1, ..., 1, 0, 0, ...0]'$ and $\boldsymbol{\xi} = [\xi_1, \xi_2, ..., \xi_n]'$.

For the symmetric 13 term Henderson filter,

$$\xi_k = 0.5 \quad \text{and} \quad k = \left\lfloor \frac{n+1}{4} \right\rfloor$$

Figure 3: Index of Italian production $\mathbf{y}$ (green), $\hat{\mathbf{m}}$ (red) and $\hat{\mathbf{m}}_k$ (blue).

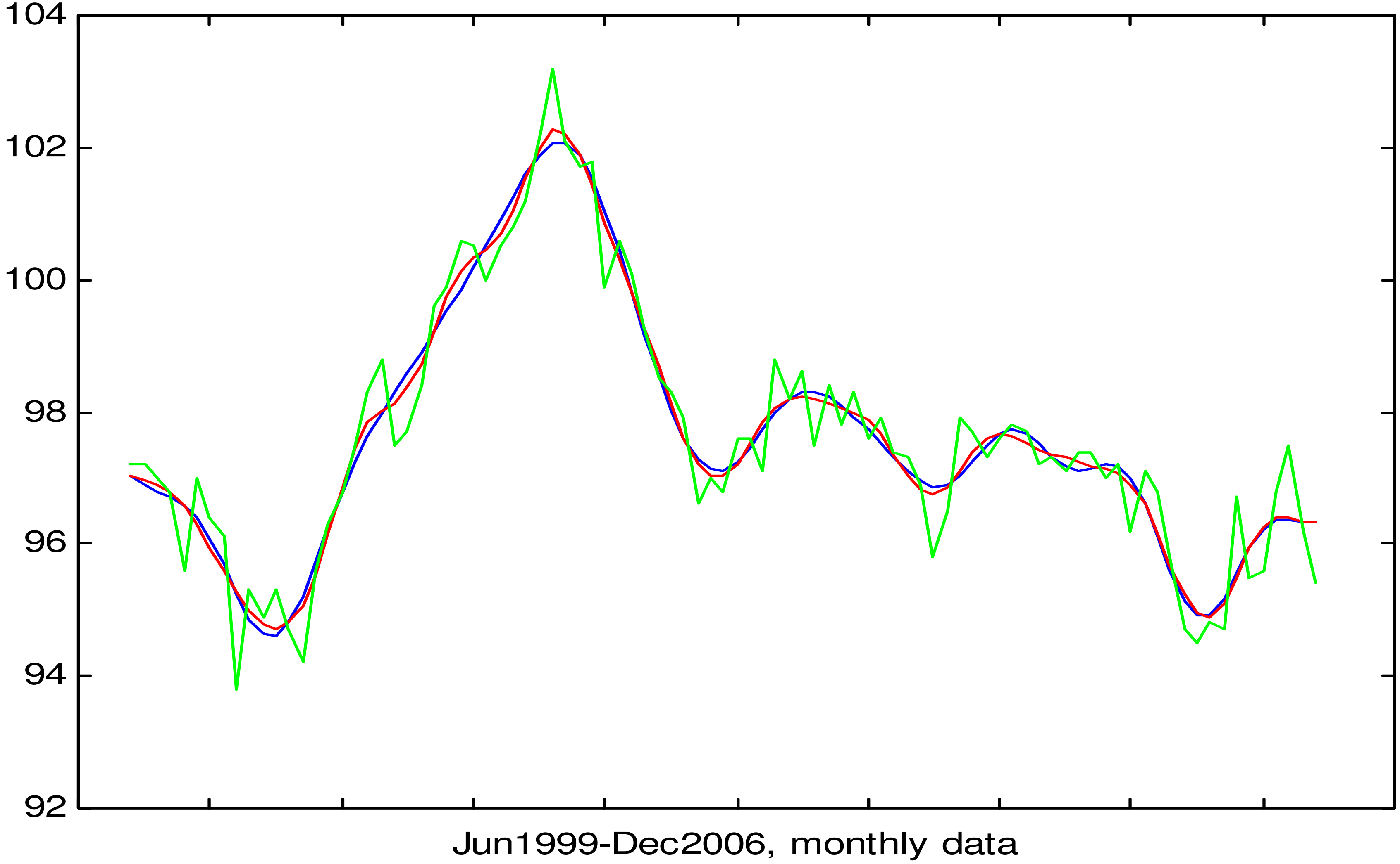
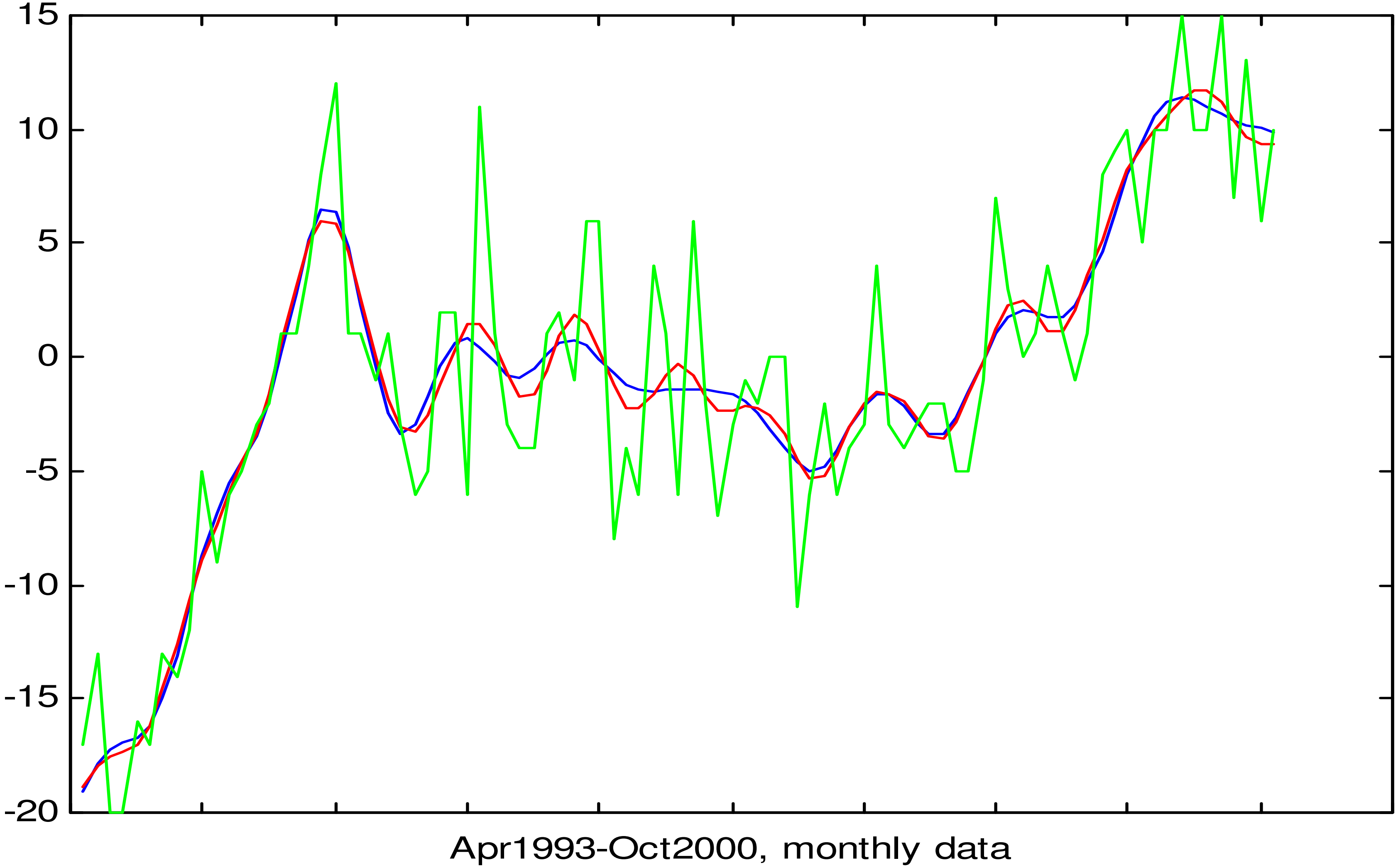


Figure 4: Euro Area Industry, Retail of Ea4 $\mathbf{y}$ (green), $\hat{\mathbf{m}}$ (red) and $\hat{\mathbf{m}}_k$ (blue).



Conclusions

Analytical results on the eigenvalues and eigenvectors of matrices associated with trend filters have been derived based on reflecting boundary conditions imposed to infinite symmetric Toeplitz banded operators.

Knowledge of the analytical form of the eigenvectors allows to represent in the time domain the periodic latent components of any time series that the filter smooths by means of the associated eigenvalues.

Inferential eigenvalue-based procedures can be developed, among the others, a strategy for a filter design in the time domain where further balancing of the trade-off between variance and bias is obtained by selecting a cut-off time or eigenvalue after which noisy components will be zero weighted.