On the Spectral Properties of Matrices Associated with Trend Filters

Luati, Alessandra and Proietti, Tommaso

Department of Statistics, University of Bologna, SEFEMEQ, University of Rome "Tor Vergata"

10 November 2008

On the Spectral Properties of Matrices Associated with Trend Filters

Alessandra Luati
Department of Statistics
University of Bologna

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

Abstract

This note is concerned with the spectral properties of matrices associated with linear smoothers. We derive analytical results on the eigenvalues and eigenvectors of smoothing matrices by interpreting the latter as perturbations of matrices belonging to algebras with known spectral properties, such as the Circulant and the generalised Tau. These results are used to characterise the properties of a smoother in terms of an approximate eigen-decomposition of the associated smoothing matrix.

Keywords Signal extraction; Smoothing; Boundary conditions; Matrix algebras.

JEL codes: C22.

1 Introduction and motivations

This note is concerned with linear smoothers that provide the estimator of a signal, $\hat{\mathbf{y}}$, as linear combinations of the observations:

$$\hat{\mathbf{y}} = \mathbf{S}\mathbf{y}. \tag{1}$$

Here $\mathbf{S}$ is the $n \times n$ smoothing matrix associated with the filter and $\mathbf{y}$ is an $n \times 1$ vector of observed values.

The rows of $\mathbf{S}$ define the equivalent kernel of the smoother and arise from a number of both parametric and nonparametric approaches: (i) local polynomial regression (see Fan and Gjbels, 1996); (ii) filtering with low-pass filters designed in the frequency domain (see for instance Baxter and King, 1999, and Christiano and Fitzgerald, 2003); (iii) wavelet multiresolution analysis (Percival and Walden, 2000); (iv) penalized least squares (Green and Silverman, 1994); (v) linear mixed models using parametric representations for the signal (Whittle, 1983).

The eigen-decomposition of $\mathbf{S}$ provides a useful characterisation of the properties of a smoother; see Buja, Hastie and Tibshirani (1989), Hastie and Tibshirani (1990) and Ruppert, Wand and Carroll, (2003). In the symmetric case, if $\mathbf{S} = \sum_{i=1}^{n} \lambda_i \mathbf{v}_i \mathbf{v}_i'$ is the spectral decomposition of the smoothing matrix, where λ_i are the ordered eigenvalues and $\mathbf{v}_i$ the corresponding eigenvectors, we can meaningfully decompose the fit as $\hat{\mathbf{y}} = \sum_{i=1}^{n} \alpha_i \lambda_i \mathbf{v}_i$, where the eigenvectors $\mathbf{v}_i$ illustrate what sequences are preserved or compressed via a scalar multiplication and α_i are the specific coefficients of the projection of $\mathbf{y}$ onto the space spanned by the eigenvectors $\mathbf{v}_i$, $\mathbf{y} = \sum_{i=1}^{n} \alpha_i \mathbf{v}_i$.

Moreover, $\text{tr}(\mathbf{S}) = \sum_{i=1}^{n} \lambda_i$ provides the number of degrees of freedom of a smoother, which is a measure of the equivalent number of parameters used to obtain the fit $\hat{\mathbf{y}}$ that allows to compare alternative filters according to their degree of smoothing. A related notion is that of the rank of a smoother.

The eigen-decomposition of a smoothing matrix is most informative if the matrix $\mathbf{S}$ is symmetric. In fact, when this is not the case, the eigenvalues and eigenvectors are complex and the interpretation of the spectral decomposition is not direct. In the nonsymmetric case Buja, Hastie and Tibshirani (1989) propose to analyse of the singular value decomposition of $\mathbf{S}$, since the singular values are always real as they represent the squared root of the eigenvalues of the symmetric $\mathbf{S}\mathbf{S}'$. Nevertheless, the right eigenvectors differ from the left eigenvectors and it is no longer clear what components are passed through by the filter or compressed.

Symmetric smoothers arise in the context of spline smoothing and from optimal signal extraction for certain classes of parametric linear mixed models (see e.g. Whittle, 1983). A relevant case in macroeconomics is the Leser-Hodrick-Prescott filter (see Leser, 1951, Hodrick and Prescott, 1997). However, nonsymmetric smoothing matrices arise in a variety of important contexts as in local polynomial regression, and more generally, when a finite impulse response (FIR) filter is designed according to some constructive principle. A common characteristic of the approaches leading to FIR filters is that a constructive principle (e.g. band–pass filtering, Baxter and King, 1999, Percival and Walden, 2000, or local polynomial reproduction, Fan and Gjibels, 1996, Cleveland and Loader, 1996) yields a two–sided filter for the central observations, using a specified bandwidth. The filter is later adapted to the boundaries and a large literature has been devoted to the estimation of the signal at the boundaries of the parameter space. The smoothing weights at the boundary are derived according to some approximation criterion, e.g. truncation, followed by normalization, or extension of the sequence according to some criterion, such as zero padding, ARIMA forecasts, etc. (see Proietti and Luati, 2009, for local polynomial regression and Christiano and Fitzgerald, 2003, for the band-pass filter). All these strategies produce a non symmetric smoothing matrix $\mathbf{S}$.

In all these instances the structure of $\mathbf{S}$ is the following (see Dagum and Luati, 2004):

$$\mathbf{S} = \begin{bmatrix} \mathbf{S}^{a}_{(m\times 2m)} & & \mathbf{O}_{(m\times n-2m)} \\ & \mathbf{S}^{s}_{(n-2m\times n)} & \\ \mathbf{O}_{(m\times n-2m)} & & \mathbf{S}^{a*}_{(m\times 2m)} \end{bmatrix} \tag{2}$$

where $\mathbf{S}^s$ is the submatrix whose rows are the symmetric filters, while $\mathbf{S}^a$ and $\mathbf{S}^{a*}$ contain the asymmetric filters to be applied to the first and last observations, respectively; the number into parentheses indicate the dimension of the submatrices, where m is the (half) bandwidth of the filter (e.g. for the Baxter and King filter, m is three years of quarterly or monthly data). Hence, the smoothing matrix is centrosymmetric, but not symmetric, with the consequence that their eigenvalues and eigenvectors are complex. Moreover, very little is known about the analytical form of such quantities, except that eigenvectors are either symmetric or skew symmetric (Weaver, 1985).

This note analyses the spectral properties of the matrices associated with linear smoothers in the case when the smoothing matrices are non-symmetric. These matrices can be interpreted as finite approximations of infinite symmetric banded Toeplitz (SBT) operators. The latter have been extensively explored, but their finite counterparts subject to boundary conditions are much more difficult to analyse (see Böttcher and Grudsky, 2005; see also Gray, 2006). The availability of eigenvalues and eigenvectors in analytical form has many desirable implications. In fact, the eigenvectors of the local polynomial regression matrices can be interpreted as the latent components of any time series that the filter smooths through the corresponding eigenvalues. Hence, eigenvalue-based inferential procedures can be developed.

2 Main results and discussion

In the ideal case of a doubly infinite sample, the matrix $\mathbf{S}$ is a SBT operator whose non null elements are the Fourier coefficients of the transfer function of the symmetric filter, $H(\nu) = \sum_{j=-h}^{h} \mathrm{w}_j e^{\imath\nu j}$, evaluated at the frequency ν, 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 $\lambda_1 \leq \max H(\nu), \lambda_n \geq \min H(\nu)$ (Grenander and Szegö, 1958). The fundamental eigenvalue distribution theorem states that when $n \to \infty$ the spectrum of $\mathbf{S}$ is dense on the set of values assumed by the transfer function of the symmetric filter.

In finite dimension, the analytical form of eigenvalues and eigenvectors is known only for few classes of matrices, which are the tridiagonal SBT and matrices belonging to some algebras, namely the Circulant, the Hartley and the generalised Tau. All these matrix algebras are associated with discrete transforms such as, respectively, the Fourier, the Hartley and the various versions of the Sine or Cosine; see, respectively, Davis (1979), Bini and Favati (1993), Bozzo and Di Fiore (1995) and the survey paper by Kailath and Sayed (1995).

By 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 the eigenvalues of $\mathbf{S}$ can be derived. The size of the perturbation depends on the matrix algebra and on the boundary conditions.

In our setting, appropriate choices are the Circulant algebra and the so-called Cosine I version of the Tau algebra (see below), that assume respectively a circular and a reflecting behavior of the series at the end (and at the beginning) of the sample. Our results will be based on the Tau algebra, but the methods apply to any of the above mentioned class of matrices. The Tau algebra has interesting properties that will be discussed in the following section, also in comparison with those of the Circulant algebra, more popular among statisticians and econometricians (Pollock, 2002).

2.1 Reflecting boundary conditions

Besides the class of circulant matrices, another class of matrices with known spectral properties in finite dimension is the $\tau_{\psi\varphi}$ algebra (Bozzo and Di Fiore, 1995), that is associated with different versions of the Sine and Cosine transforms and constitutes a generalisation of the τ family (Bini and Capovani, 1983). An $n \times n$ matrix $\mathbf{H}$ belongs to the $\tau_{\psi\varphi}$ class if and only if

$$\mathbf{T}_{\psi\varphi}\mathbf{H} = \mathbf{H}\mathbf{T}_{\psi\varphi},$$

where

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

and $\psi, \varphi = 0, 1, -1$. The elements h_{ij} of the matrices in $\tau_{\psi\varphi}$ satisfy the cross sum property $\mathrm{h}_{i-1,j} + \mathrm{h}_{i+1,j} = \mathrm{h}_{i,j-1} + \mathrm{h}_{i,j+1}$ subject to boundary conditions determined by ψ and φ. For the original τ algebra arising when $\psi = \varphi = 0$ the boundary conditions are $\mathrm{h}_{0j} = \mathrm{h}_{i0} = \mathrm{h}_{n+1,j} = \mathrm{h}_{i,n+1} = 0$, $i, j = 1, ..., n$ and all the matrices in τ can be then derived given their first row elements. Still based on the first row of $\mathbf{H}$ but more appropriate for our purposes, since it allows to obtain the eigenvalues and eigenvectors of $\mathbf{H} \in \tau_{\psi\varphi}$ in an amenable form, is the following way to construct $\mathbf{H}$ as a linear combination of powers of $\mathbf{T}_{\psi\varphi}$ (see Bini and Capovani, 1983, Proposition 2.2).

Let $\mathbf{h}' = [\mathrm{h}_{11}, \mathrm{h}_{12}, ..., \mathrm{h}_{1n}]$ be the first row of $\mathbf{H}$. Then

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

where $\mathbf{c}$ is the solution of the upper triangular system $\mathbf{Q}\mathbf{c} = \mathbf{h}$ and $\mathbf{Q}$ is the matrix whose j-th column equals the first column of $\mathbf{T}_{\psi\varphi}^{j-1}$. It follows that the eigenvalues of $\mathbf{H}$ are given by

$$\xi_i = \sum_{j=1}^{n} \vartheta_i^{j-1} c_j \tag{3}$$

where $\vartheta_i, i = 1, .., n$, are the eigenvalues of $\mathbf{T}_{\psi\varphi}$. The eigenvectors of $\mathbf{H}$ are the same of $\mathbf{T}_{\psi\varphi}$.

Let us consider the reflecting hypothesis such 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 $2m+1$-term estimator corresponds to the real time filter $\{\mathrm{w}_m, \mathrm{w}_{m-1} + \mathrm{w}_m, ..., \mathrm{w}_1 + \mathrm{w}_2, \mathrm{w}_0 + \mathrm{w}_1\}$, made of $m + 1$ terms. With the constraint of being centrosymmetric, the reflecting matrix $\mathbf{H}$ belongs to the τ_{11} algebra and its first row is the vector

$$\mathbf{h}' = [\mathrm{w}_0 + \mathrm{w}_1, \mathrm{w}_1 + \mathrm{w}_2, \mathrm{w}_2 + \mathrm{w}_3, ..., \mathrm{w}_{m-1} + \mathrm{w}_m, \mathrm{w}_m, 0, ..., 0]\,. \tag{4}$$

With these premises, we are able to construct $\mathbf{H} \in \tau_{11}$ associated with the symmetric filter $\{\mathrm{w}_{-m}, ..., \mathrm{w}_0, \mathrm{w}_1, ..., \mathrm{w}_m\}$. Given $\mathbf{H}$, we will denote its spectrum by $\sigma(\mathbf{H})$ and its 2-norm by $\|\mathbf{H}\|_2 = \sqrt{\rho(\mathbf{H}'\mathbf{H})}$ where $\rho(\mathbf{H})$ is the spectral radius of $\mathbf{H}$, which is the maximum modulus of its eigenvalues. With this notation, we may state the following result where, for sake of notation, we use the Pochhammer symbol for rising factorial, $(j)_q = j(j+1)(j+2)...(j+q-1)$, for $q = 0, \ldots, \left\lfloor \frac{m-j-1}{2} \right\rfloor$, the latter term denoting the largest integer less than or equal to $\frac{m-j-1}{2}$, and $(j)_q = 1$ for $q = 0$.

Theorem 1 *Let* $\mathbf{S}$ *and* $\mathbf{H}$ *be* $n \times n$ *smoothing matrices associated with the symmetric filter* $\{\mathrm{w}_{-m}, ..., \mathrm{w}_0, ..., \mathrm{w}_m\}$*, and let* $\mathbf{H} \in \tau_{11}$*. Hence,* $\forall \lambda \in \sigma(\mathbf{S})$*,* $\exists i \in \{1, 2, .., n\}$ *such that*

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

where

$$\xi_i = \sum_{j=1}^{m+1} \left(2\cos\frac{(i-1)\pi}{n}\right)^{j-1} \left[\mathrm{w}_{j-1} + \sum_{q=0}^{\left\lfloor \frac{m-j-1}{2} \right\rfloor} \frac{(-1)^{q+1}(j)_q}{(q+1)!}(j+2q+1)\mathrm{w}_{j+2q+1}\right] \quad (5)$$

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

The proof is in the appendix. As a by-product, theorem 1 gives the eigenvalues of $\mathbf{H} \in \tau_{11}$, with first row equal to (4), as an explicit function of the filter weights, as shown in (5). The corresponding eigenvectors are (Bozzo and Di Fiore, 1995):

$$\mathbf{z}_i = k_i \left[\cos\frac{(2j-1)(i-1)\pi}{2n}\right]_j, j = 1, 2, ..., n \quad (6)$$

with $k_i = \frac{1}{\sqrt{2}}$ for $i = 1$ and $k_i = 1$ for $i > 1$.

Theorem 1 provides an upper bound to the size of the perturbation of the eigenvalues of $\mathbf{S}$ with respect to those of $\mathbf{H}$, for which an exact analytical expression is available. The quantity δ_H measures how much the eigenvalues of a smoothing matrix move away from the eigenvalue distribution of the corresponding matrix in τ_{11}. The eigenvalue distribution of $\mathbf{H}$ can be visualised as the plot of the eigenvalues (5) against n and provides a discrete approximation to the transfer function of the symmetric filter. What follows is that δ_H can be chosen as a measure of how much the absolute eigenvalues of $\mathbf{S}$ deviate from the gain function of the associated filter.

We now discuss the advantages of assuming reflecting rather than circular boundary conditions. First, all the operators belonging to τ algebras have real eigenvalues and eigenvectors. All the computations related to this class can therefore be done in real arithmetic. Another important aspect is that in general Circulant-to-Toeplitz corrections produce perturbations that are not smaller than Tau-to-Toeplitz corrections, since while $\mathbf{H}$ is structured as (2), a circulant matrix has nonzero corrections in the top right and bottom left $m \times m$ blocks. When the elements of the central-block matrix are the same, this results in a greater perturbation. Finally, $\mathbf{H}$ has n distinct eigenvalues compared to the at most $\frac{n-1}{2}+1$ of a circulant matrix and so $\sigma(\mathbf{H})$ provides a better approximation to $H(\nu)$, $\nu \in (0, \pi)$.

We now consider the eigenvectors. In general, the analytical expression of the eigenvectors of a smoothing matrix cannot be derived using the perturbation theory, not even in an approximate form. However, evaluating the action of $\mathbf{S}$ on the eigenvectors of $\mathbf{H}$, we are able to show that, unless for the boundaries, the latent components of $\mathbf{S}$ can be fairly approximated by those of $\mathbf{H}$. In fact, let us decompose the time series $\mathbf{y}$ as a linear combination of the n known real and orthogonal latent components represented by the eigenvectors of $\mathbf{H}$, $\mathbf{y} = \theta_1\mathbf{z}_1 + \theta_2\mathbf{z}_2 + ... + \theta_n\mathbf{z}_n$ where the $\mathbf{z}_i$ are given by (6) and $\boldsymbol{\theta} = [\theta_1, ..., \theta_n]'$ is a vector of coefficients. It follows from theorem 1 that

$$\mathbf{S}\mathbf{y} = \sum_{i=1}^{n} \theta_i \xi_i \mathbf{z}_i + \sum_{i=1}^{n} \theta_i \boldsymbol{\Delta}_{\mathbf{H}} \mathbf{z}_i \tag{7}$$

where $\boldsymbol{\Delta}_{\mathbf{H}}\mathbf{z}_i$ is a vector of zeros except for the first and last m coordinates, i.e. $\boldsymbol{\Delta}_{\mathbf{H}}\mathbf{z}_i = \begin{bmatrix} \mathbf{z}_i^{*'} & \mathbf{0}' & \mathbf{E}_h\mathbf{z}_i^{*'} \end{bmatrix}'$ and $\mathbf{z}_i^* = \sum_{j=1}^{q}(\mathbf{S}_{ij} - \mathbf{H}_{ij})z_{ij}$ for $q = m+1, ..., 2m$ and $i = 1, 2, ..., m$. Due to the fact that the elements of both $\mathbf{S}$ and $\mathbf{H}$ add up to one and their absolute values are in general smaller than one, the values in $\mathbf{z}_i^*$ and in $\mathbf{E}_h\mathbf{z}_i^*$ are almost zero. This holds for all $n > 2m$.

As a consequence of (7), the eigenvectors of $\mathbf{H}$ can be interpreted as the periodic latent components of any time series, that the filter modifies through multiplication by the corresponding eigenvalues. Specifically, by (5) and (6), (7) can be written as

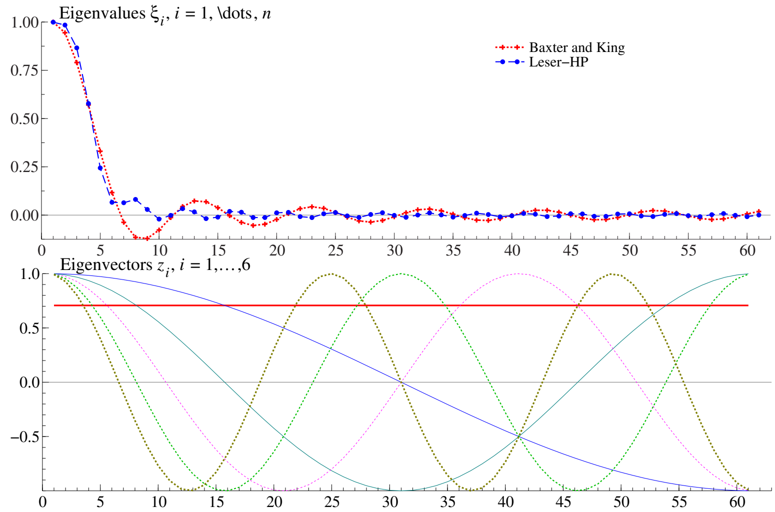
$$\mathbf{S}\mathbf{y} = \sum_{i=1}^{k} \theta_i \xi_i \mathbf{z}_i + \sum_{i=k+1}^{n} \theta_i \xi_i \mathbf{z}_i + \sum_{i=1}^{n} \theta_i \boldsymbol{\Delta}_{\mathbf{H}} \mathbf{z}_i,$$

i.e. the series $\mathbf{y}$ can be decomposed as the sum of k long-period components that the filter leaves unchanged or smoothly shrinks, and these account for the signal, and $n-k$ high-frequency components that will be almost suppressed, and these account for the noise. The choice of k turns out to be a filter design problem in the time domain. There is a mathematically elegant exact solution, which occurs if rank$(\mathbf{H}) = k$ that is $\hat{\mathbf{m}}$ belongs to the column space $\mathcal{C}(\mathbf{H})$ and $\boldsymbol{\varepsilon}$ lies in the null space $\mathcal{N}(\mathbf{H})$. In practice, even if many of the eigenvalues are close to zero, $\mathbf{H}$ is full rank and therefore we may only look for an approximate solution that consists of choosing a cut-off time or a cut-off eigenvalue.

3 Applications

The results of the previous section can be used to provide the eigen-decomposition of the smoothing matrix corresponding to the low-pass Baxter and King filter with cutoff frequency corresponding to 10 years of quarterly data, and to compare it to the Leser-Hodrick-Prescott filter for quarterly data (smoothing parameter 1600). The eigenvalues ξ_i are reproduced in the top panel of figure 1, which are constructed for $n = 61$. The plot reveals that these are very similar and that only the first six are relevant to describe the properties of the filters. The corresponding eigenvectors are plotted in the bottom panel. It is also clear from the plot that the Leser-Hodrick-Prescott component suffers less from the leakage from periodic features z_i with smaller period.

Figure 1: Eigen-decomposition of the smoothing matrices corresponding to the Baxter and King low-pass filter with cutoff frequency corresponding to 10 years (quarterly data) and to the Leser-Hodrick-Prescott filter with smoothing parameter 1600.



The results of the preceding sections can also be relevant for the desing of a filter in the time domain. The method consists of modifying $\mathbf{S}$ so that $n-k$ high frequency noisy components receive zero weight. This is done through the spectral decomposition of $\mathbf{H}$.

Decomposing $\mathbf{S} = \mathbf{H} + \mathbf{\Delta_H}$ and $\mathbf{H} = \mathbf{Z}\mathcal{X}\mathbf{Z}'$, where $\mathcal{X} = \text{diag}\{\xi_1, \xi_2, ..., \xi_n\}$, and writing $\mathbf{y} = \mathbf{Z}\boldsymbol{\theta}$, we get

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

where $\mathcal{X}_k$ is the matrix obtained by replacing with zeros the eigenvalues of $\mathbf{H}$ that are smaller than a cut-off eigenvalue ξ_k and $\mathbf{\Delta_H}\mathbf{Z}\boldsymbol{\theta}$ is a null vector except for the first and last elements that account for the boundary conditions. Turning to the original coordinate system and arranging the boundaries, we get the new estimator

$$\begin{aligned}\mathbf{S}_k &= \mathbf{H}_k + \mathbf{\Delta}_k + \mathbf{\Delta_H} \\ &= \mathbf{H}_{(k)} + \mathbf{\Delta_H}\end{aligned}$$

where $\mathbf{H}_{(k)}$ is the matrix with boundaries equal to those of $\mathbf{H}$ and interior equal to that of $\mathbf{H}_k = \mathbf{Z}\mathcal{X}_k\mathbf{Z}'$. In other words, $\mathbf{H}_{(k)}$ is structured like (2) with $\mathbf{H}^a_k = \mathbf{H}^a$, $\mathbf{H}^{a*}_k = \mathbf{E}_h\mathbf{H}^a\mathbf{E}_h$ and $\mathbf{H}^s_k = [\mathbf{Z}\mathcal{X}_k\mathbf{Z}']^s$. Hence a new smoothing matrix is obtained, $\mathbf{S}_k$, and consequently new trend estimates, say $\hat{\mathbf{m}}_k$.

In practice, the procedure is very easy to apply. In fact, given a symmetric filter, it consists of: obtaining $\mathbf{H}$, replacing it by $\mathbf{H}_k$ and then adjusting the boundaries with suitable chosen asymmetric filters to get $\mathbf{S}_k$.

4 Appendix: proof of Theorem 1

Let us write $\mathbf{S} = \mathbf{H} + \mathbf{\Delta_H}$. The matrix $\mathbf{H}$ is diagonalised by the orthogonal matrix

$$\mathbf{Z} = \sqrt{\frac{2}{n}}\left[k_j \cos\frac{(2i-1)(j-1)\pi}{2n}\right]_{ij}, i, j = 1, 2, ..., n$$

where $k_j = \frac{1}{\sqrt{2}}$ for $j = 1$ and $k_j = 1$ for $j > 1$ which satisfies $\|\mathbf{Z}\|_2\|\mathbf{Z}^{-1}\|_2 = 1$. The spectrum of $\mathbf{H}$ is $\sigma(\mathbf{H}) = \{\xi_1, \xi_2, ..., \xi_n\}$, where

$$\xi_i = \sum_{j=1}^{n}\left(2\cos\frac{(i-1)\pi}{n}\right)^{j-1} c_j$$

which follows by (3) and by the fact that the eigenvalues of $\mathbf{T}_{11}$ are (Bini and Capovani, 1983)

$$\vartheta_i = 2\cos\frac{(i-1)\pi}{n}.$$

Setting $\delta_H = \|\mathbf{\Delta_H}\|_2$ and applying the Bauer-Fike theorem (Bauer and Fike, 1960) with the 2-norm as an absolute norm gives

$$\left|\lambda - \sum_{j=1}^{n}\left(2\cos\frac{(i-1)\pi}{n}\right)^{j-1} c_j\right| \le \delta_H.$$

We now prove that $c_j = 0$ for $j > m+1$, so that the above summation involves just $m+1$ terms instead of n. It follows by the Cramer rule that, explicitly,

$$c_j = \frac{\det \mathbf{Q}\,[j, \mathbf{h}]}{\det \mathbf{Q}}$$

where $\mathbf{Q}\,[j, \mathbf{h}]$ is the matrix obtained replacing the j-th column of $\mathbf{Q}$ by the vector $\mathbf{h}$. The matrix $\mathbf{Q}$ is upper triangular with ones on the diagonal so its its determinant is equal to one and since the generic element h_j of $\mathbf{h}$ is null for $j > m+1$ it follows that $\det \mathbf{Q}\,[j, \mathbf{h}] = 0$ and c_j will be null as well.

Finally, we prove that

$$c_j = \mathrm{w}_{j-1} + \sum_{q=0}^{\left\lfloor\frac{m-j-1}{2}\right\rfloor} \frac{(-1)^{q+1}(j)_q}{(q+1)!}(j+2q+1)\mathrm{w}_{j+2q+1}. \tag{8}$$

This expression can be directly verified by calculating $\det \mathbf{Q}\,[j, \mathbf{h}]$ for all j. Here in the following, we prove it by induction over $j = 1, ..., m+1$, with $m \in \mathbb{N}$.

- For $j = 1$, $c_1 = \mathrm{w}_0 + \sum_{q=0}^{\frac{m-2}{2}}(-1)^{q+1}2\mathrm{w}_{2q+2}$ which follows by $(1)_q = q!$ and by simple algebra. The linear system $\mathbf{Qc} = \mathbf{h}$ can be written as $\mathbf{c} = \mathbf{Q}^{-1}(\mathbf{h}_1 + \mathbf{h}_2)$ with $\mathbf{h}_1 = [\mathrm{w}_0, \mathrm{w}_1, ..., \mathrm{w}_m, 0, ..., 0]'$ and $\mathbf{h}_2 = [\mathrm{w}_1, \mathrm{w}_2, ..., \mathrm{w}_m, 0, ..., 0]'$, both n-dimensional vectors. Since the first row of $\mathbf{Q}^{-1}$ is the vector $[1, -1, -1, 1, 1, 1, -1, -1, ...]$ we have that $c_1 = (\mathrm{w}_0 + \mathrm{w}_1) - (\mathrm{w}_1 + \mathrm{w}_2) - (\mathrm{w}_2 + \mathrm{w}_3) + (\mathrm{w}_2 + \mathrm{w}_4) + ... + (-1)^{\left\lfloor\frac{m-2}{2}\right\rfloor+1}2\mathrm{w}_{2\left\lfloor\frac{m-2}{2}\right\rfloor+2}$ and therefore (8) holds for $j = 1$.

- For $j = m$, $c_m = \mathrm{w}_{m-1}$ as it is immediate to see given that the summation in q was defined for non negative values of $\frac{m-j-1}{2}$. All the more so, it implies that $c_{m+1} = \mathrm{w}_m$. Hence we have showed that (8) holds for $j = 1$ and that, if it holds for $j = m$ then it holds for $j = m + 1$. This proves that (8) is true for all $m \in \mathbb{N}$. The proof of theorem 2 is therefore complete ■

Acknowledgements

We would like to kindly thank Dario Bini for the valuable suggestions and discussions.

References

Bauer F., Fike C. (1960), Norms and Exclusion Theorems, *Numerische Mathematik*, 2, 137-141.

Baxter, M., King, R.G. (1999). Measuring Business Cycles: Approximate Band-Pass Filters for Economic Time Series. *The Review of Economics and Statistics*, 81, 575593.

Bini D., Capovani M. (1983), Spectral and computational properties of band symmetric Toeplitz matrices, *Linear Algebra and its Applications*, 52/53, pp. 99-126 .

Bini D., Favati P. (1993), On a matrix algebra related to the discrete Hartley transform. *SIMAX*, 14, 2, 500-507

Böttcher A., Grudsky, S.M. (2005), *Spectral Properties of Banded Toeplitz Matrices*, Siam.

Bozzo E., Di Fiore C. (1995), On the use of certain matrix algebras associated with discrete trigonometric transforms in matrix displacement decomposition, *Siam J. Matrix anal. Appl.*, 16, 1, 312-326.

Buja A., Hastie T.J., Tibshirani R.J.(1989), Linear Smoothers and Additive Models, *The Annals of Statistics*, 17, 2, 453-555.

Christiano L.J., Fitzgerald T.J., (2003), The band pass filter, *International Economic Review*, 44, 435-465.

Cleveland, W.S., Loader, C.L. (1996). Smoothing by Local Regression: Principles and Methods. In W. Härdle and M. G. Schimek, editors, *Statistical Theory and Computational Aspects of Smoothing*, 1049. Springer, New York.

Dagum, E.B., Luati, A. (2004). A Linear Transformation and its Properties with Special Applications in Time Series, *Linear Algebra and its Applications*, 338, 107-117.

Davis, P.J., (1979). *Circulant matrices*, Wiley, New York.

Fan J. and Gjibels I. (1996). *Local Polynomial Modelling and its Applications*, Chapman and Hall, New York.

Gray R.M. (2006) *Toeplitz and Circulant Matrices: A review*, Foundations and Trends in Communications and Information Theory, Vol 2, Issue 3, pp 155-239.

Green P.J. and Silverman, B.V. (1994) *Nonparametric Regression and Generalized Linear Models: a Roughness Penalty Approach*. Chapman & Hall, London.

Grenander, U., Szegö, G. (1958). *Toeplitz Forms and Their Applications*, University of California Press.

Hastie T.J. and Tibshirani R.J.(1990), *Generalized Additive Models*, Chapman and Hall, London.

Hodrick, R., and Prescott, E.C. (1997). Postwar U.S. Business Cycle: an Empirical Investigation, *Journal of Money, Credit and Banking*, 29, 1, 1-16.

Kailath T., Sayed A. (1995) Displacement Structure: Theory and Applications, *SIAM Review*, 37, 3, 297-386.

Leser, C.E.V. (1961). A Simple Method of Trend Construction, *Journal of the Royal Statistical Society B*, 23, 91-107.

Percival, D.B., Walden, A.T. (2000). *Wavelet Methods for Time Series Analysis*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, Cambridge.

Pollock, D.S.G. (2002), Circulant matrices and time-series analysis, *International Journal of Mathematical Education in Science and Technology*, 33, 2, 1, 213-230(18)

Proietti, T., Luati, A. (2009). Real Time Estimation in Local Polynomial Regression, with an Application to Trend-Cycle Analysis, *Annals of Applied Statistics*, forthcoming.

Ruppert D., Wand, M.J., Carroll R.J. (2003). *Semiparametric regression*, Cambridge University Press.

Weaver J.R. (1985), Centrosymmetric (Cross-Symmetric) Matrices, their Basic Properties, Eigenvalues, Eigenvectors, *Amer. Math. Monthly*, 92, 711-717.

Whittle P. (1983). *Prediction and Regulation by Linear Least Squares Methods*, Second edition. Basil Blackwell, Oxford.