Estela Bee Dagum, Alessandra Luati

A Linear Transformation and its Properties with Special Applications in Time Series Filtering

Serie Ricerche 2002, n.2

Dipartimento di Scienze Statistiche "Paolo Fortunati"
Università degli studi di Bologna

Finito di stampare nel mese di Febbraio 2002
presso le Officine Grafiche Tecnoprint
Via del Legatore 3, Bologna

1 Introduction

In time series analysis, it is often assumed that the data generating process can be decomposed into various unobservable components representing the trend, cyclical fluctuations, seasonal effects and irregulars. These latent variables are estimated by applying linear filters or systems of weights to the observations, in a moving manner. The filters can be arranged in matrix form such that, applied to the vector of observations, produce the corresponding estimated values. If the linear filters are symmetric, say of length $2m+1$, with m positive integer, and applied to a series of length $N > 2m+1$, then it is evident that the components cannot be estimated for the first and last m observations. However, since for policy and decision making is of great importance to have estimates of the latent variables up to and including the most recent observations, asymmetric filters must be applied to the beginning and ending m values of the series. The entire predictor matrix is here shown to be invariant with respect to a linear transformation called t and which results from pre- and post-multiplication of a given matrix by two permutation matrices of suitable dimensions. In particular, we show that the predictor matrix is centrosymmetric and that it is formed by a submatrix of symmetric weights (to be applied to central observations) which is t-invariant or, equivalently, rectangular centrosymmetric, and by submatrices of asymmetric weights (to be applied to initial and final observations) which are the t-transform of each other.

We would like to remark that the t-transformation has been improperly (see, for example, Farebrother, [7]) referred to either as a 'reflection' by Weaver [23] or as a 'rotation' by Krafft and Schaever [11]. In this paper, we define and study the properties of the t-transformation and highlight its role in time series filtering.

Section 2 introduces the t-transformation, and derives its basic properties; section 3 deals with the problem of time series linear filtering by means of the smoothing matrices that represent the transformations acting on the data to produce smooth estimates of the latent variables. Finally, section 4 provides examples of the role of the t-transformation on most often applied trend smoothers.

2 The t-transformation and centrosymmetric matrices

Let $\mathbb{R}^{m\times n}$ denote the set of $m\times n$ real matrices and let $\mathbf{A}\in\mathbb{R}^{m\times n}$ be the matrix of generic element $a_{ij}, i=1,\dots,m,\ j=1,\dots,n$ with $m,n<\infty$. The t-transformation is defined as follows:

$$\begin{aligned} t &: \mathbb{R}^{m\times n}\longrightarrow\mathbb{R}^{m\times n} \\ \mathbf{A} &\mapsto t(\mathbf{A}) \end{aligned}$$

such that

$$a_{ij}\longmapsto a_{m+1-i,n+1-j} \qquad (1)$$

for $i=1,\dots,m$ and $j=1,\dots,n$.

The action of t on $\mathbf{A}$ can be described by means of two permutation matrices of equal form but different dimensions. Precisely, if $\mathbf{A}\in\mathbb{R}^{m\times n}$, then

$$t(\mathbf{A})=\mathbf{E}_m\mathbf{A}\mathbf{E}_n \qquad (2)$$

where $\mathbf{E}_k\in\mathbb{R}^{k\times k}$ is the permutation matrix with ones on the cross diagonal (bottom left to top right) and zeros elsewhere. In other words, t acts on $\mathbf{A}$ reversing the order of its rows by premultiplication for $\mathbf{E}_m$ and then reversing the order of its columns by postmultiplication for $\mathbf{E}_n$.

Definition (Weaver [23]). If $\mathbf{A}\in\mathbb{R}^{m\times n}$ and $a_{ij}=a_{m+1-i,n+1-j}$, for all $i=1,\dots,m$ and $j=1,\dots,n$ then $\mathbf{A}$ is *rectangular centrosymmetric*.

Centrosymmetric matrices can be found in many applications in statistics and time series analysis. The most commonly known are: the permutation matrices like $\mathbf{E}_k$, used, among others, for the computational solution of various linear model estimation problems (see Kontoghiorghes [10]); the symmetric Toeplitz matrices $\mathbf{R}\in\mathbb{R}^{m\times m}$ of generic element $r_{ij}=r_{i+k,j+k}=r_{ji}, i,j=1,\dots,m, k=1,\dots,m-1$ for the autocorrelation of stationary time series (see Trench [18], [19] and [20]); the commutation matrix $\mathbf{K}_{mn}\in\mathbb{R}^{mn\times mn}$ such that $\mathbf{K}_{mn}vec\mathbf{A}=\mathbf{K}_{mn}vec\mathbf{A}^T$, where $vec\mathbf{A}$ is the vector obtained by stacking the column of the matrix $\mathbf{A}$ one underneath the other (see Magnus and Neudecker [14]). Furthermore, Iosifescu's [12] and Kimura's [9] transition matrices for some Markov chain in genetic problems are centrosymmetric. Recently, (rectangular) centrosymmetric matrices also can be viewed as a particular case of (generalized) reflexive matrices whose properties have been recently employed in linear least-squares problems (Chen [3]). Concerning the properties of centrosymmetric matrices, useful references can be found in Andrew [1].

2.1 Properties of the t-transformation

The t-transformation inherits desirable properties from the properties of the permutation matrix $\mathbf{E}_k$ which is: (a) symmetric (b) orthogonal and (c) a reflection, *i.e.*

$$\mathbf{E}_k \overset{(a)}{=} \mathbf{E}_k^T \overset{(b)}{=} \mathbf{E}_k^{-1} \text{ and } \mathbf{E}_k^2 \overset{(c)}{=} \mathbf{I}_k$$

where $\mathbf{I}_k$ is the identity matrix of order k. These (basic) properties are:

1. t is linear.
2. $\mathbf{A} \in \mathbb{R}^{m\times n}, t(t(\mathbf{A})) = \mathbf{A}$.
3. $\mathbf{A} \in \mathbb{R}^{m\times n}, t^{-1}(\mathbf{A}) = t(\mathbf{A})$.
4. $\mathbf{A} \in \mathbb{R}^{m\times n}, t(\mathbf{A}^T) = (t(\mathbf{A}))^T$.
5. $\mathbf{A} \in \mathbb{R}^{m\times m}, t(\mathbf{A}) = \mathbf{E}_m^{-1}\mathbf{A}\mathbf{E}_m$.
6. $\mathbf{A} \in \mathbb{R}^{m\times n}, \mathbf{B} \in \mathbb{R}^{n\times p}, t(\mathbf{AB}) = t(\mathbf{A})\,t(\mathbf{B})$.
7. $\mathbf{A} \in \mathbb{R}^{m\times m}, t(\mathbf{A}^n) = [t(\mathbf{A})]^n$.
8. $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{m\times m}, \det\{\mathbf{AB}\} = \det\{t(\mathbf{A})\}\det\{t(\mathbf{B})\}$.
9. $\mathbf{A} \in \mathbb{R}^{m\times m}, rank\{\mathbf{A}\} = m, t(\mathbf{A}^{-1}) = [t(\mathbf{A})]^{-1}$.

Proofs.

1. **(a)** $t(\mathbf{A}+\mathbf{B}) = \mathbf{E}_m(\mathbf{A}+\mathbf{B})\mathbf{E}_n = \mathbf{E}_m\mathbf{A}\mathbf{E}_n + \mathbf{E}_m\mathbf{B}\mathbf{E}_n = t(\mathbf{A}) + t(\mathbf{B})$.

(b) $t(\lambda\mathbf{A}) = \mathbf{E}_m(\lambda\mathbf{A})\mathbf{E}_n = \lambda\mathbf{E}_m\mathbf{A}\mathbf{E}_n = \lambda t(\mathbf{A})$.

2. $t(t(\mathbf{A})) = \mathbf{E}_m(\mathbf{E}_m\mathbf{A}\mathbf{E}_n)\mathbf{E}_n = (\mathbf{E}_m\mathbf{E}_m)\mathbf{A}(\mathbf{E}_n\mathbf{E}_n) = \mathbf{A}$.

3. It follows by **2**, given $t^{-1}: \mathbb{R}^{m\times n} \to \mathbb{R}^{m\times n}$, $t(\mathbf{A}) \mapsto \mathbf{A}$.

4. $t(\mathbf{A}^T) = \mathbf{E}_n\mathbf{A}^T\mathbf{E}_m = \mathbf{E}_n^T\mathbf{A}^T\mathbf{E}_m^T = (\mathbf{E}_m\mathbf{A}\mathbf{E}_n)^T = (t(\mathbf{A}))^T$.

5. It follows by (a) and (b).

6. $t(\mathbf{AB}) = \mathbf{E}_m\mathbf{A}\mathbf{B}\mathbf{E}_p = \mathbf{E}_m\mathbf{A}\mathbf{E}_n\mathbf{E}_n\mathbf{B}\mathbf{E}_p = t(\mathbf{A})\,t(\mathbf{B})$.

7. The proof is by induction. For $n = 1$, it trivially holds. For $n+1$, $t(\mathbf{A}^{n+1}) = t(\mathbf{A}^n\mathbf{A}) =$ by **6** $= t(\mathbf{A}^n)\,t(\mathbf{A}) =$ for the inductive hypothesis $= [t(\mathbf{A})]^n\,t(\mathbf{A}) = [t(\mathbf{A})]^{n+1}$.

8. By the Theorem of Binet and Cauchy: $\det\{\mathbf{AB}\} = \det\{\mathbf{E}_m^{-1}\mathbf{A}\mathbf{E}_m\mathbf{E}_m^{-1}\mathbf{B}\mathbf{E}_m\} = \det\{t(\mathbf{A})\}\det\{t(\mathbf{B})\}$.

9. $t(\mathbf{A}^{-1}) = \mathbf{E}_m\mathbf{A}^{-1}\mathbf{E}_m = \mathbf{E}_m^{-1}\mathbf{A}^{-1}\mathbf{E}_m^{-1} = (\mathbf{E}_m\mathbf{A}\mathbf{E}_m)^{-1} = (t(\mathbf{A}))^{-1}$.

Here in the following, we briefly comment the most relevant of the above properties.

Coordinate systems. Property **5** states that a relation of similitude holds between a square matrix $\mathbf{A}$ and $t(\mathbf{A})$. Hence $\mathbf{A}$ and $t(\mathbf{A})$ represent the same linear transformation with respect to different bases. In particular, if $\mathbf{A}$ represents a linear transformation $f: \mathbb{R}^m \to \mathbb{R}^m, \mathbf{v} \longmapsto f(\mathbf{v})$, with respect to the canonical basis $\mathcal{E} = \{\mathbf{e}_1, \cdots, \mathbf{e}_m\}$, where $\mathbf{e}_j \in \mathbb{R}^m$, $\forall j = 1, \cdots, m$ is the m-dimensional vector with all null components except for the j-th, is chosen, then $t(\mathbf{A})$ represents the same linear transformation with respect to the basis $\mathcal{E}' = \{\mathbf{e}_m, \cdots, \mathbf{e}_1\}$. $\mathbf{E}_m$ can be interpreted as the changing basis matrix from $\mathcal{E}$ to $\mathcal{E}'$ having, as columns, the components of the vectors of the basis $\mathcal{E}'$ taken with respect to the basis $\mathcal{E}$.

Linear transformations. Concerning the action of $\mathbf{A} \in \mathbb{R}^{m\times m}$ and $t(\mathbf{A})$ on a same vector $\mathbf{v}^T = \begin{bmatrix} v_1 & v_2 & \cdots & v_m \end{bmatrix}$, we have that $\mathbf{Av} = \mathbf{w}$, where $\mathbf{w}^T = \begin{bmatrix} w_1 & w_2 & \cdots & w_m \end{bmatrix}$. On the other hand, $t(\mathbf{A})\,\mathbf{v} = \mathbf{E}_m\mathbf{A}\mathbf{E}_m\mathbf{v} = \mathbf{E}_m\mathbf{A}\mathbf{v}^* = \mathbf{E}_m\mathbf{z}^* = \mathbf{z}$, where $\mathbf{v}^{*T} = \begin{bmatrix} v_m & v_{m-1} & \cdots & v_1 \end{bmatrix}$, $\mathbf{z}^{*T} = \begin{bmatrix} z_m & z_{m-1} & \cdots & z_1 \end{bmatrix}$, and $\mathbf{z}^T = \begin{bmatrix} z_1 & z_2 & \cdots & z_m \end{bmatrix}$. Hence, the components of $\mathbf{w}$ and of $\mathbf{z}$ are the coordinates of the same vector $f(\mathbf{v})$ taken with respect to the bases $\mathcal{E}$ and $\mathcal{E}'$ respectively.

Eigenvalue and eigenvectors. Another relevant consequence of **5** is that $\mathbf{A}$ and $t(\mathbf{A})$ have the same spectrum, $\sigma(\mathbf{A})$. Notice that if $\mathcal{A}_\lambda$ is the eigenspace of $\mathbf{A}$ corresponding to the eigenvalue λ, then $t(\mathcal{A}_\lambda)$ is the eigenspace of $t(\mathbf{A})$ corresponding to the same eigenvalue and it coincides with the subspace of $\mathbb{R}^m$ spanned by the transformed of the eigenvectors spanning $\mathcal{A}$, *i.e.* if $\mathcal{A}_\lambda = \{v \in \mathbb{R}^m, \mathbf{A}v = \lambda v\}$ then $t(\mathcal{A}_\lambda) = \{t(v) \in \mathbb{R}^m, t(\mathbf{A})\,t(v) = \lambda t(v)\}$. In fact, $\mathbf{A}v = \lambda v \Rightarrow \mathbf{E}_m\mathbf{A}\mathbf{E}_m\mathbf{E}_m v\mathbf{E}_1 = \lambda\mathbf{E}_m v\mathbf{E}_1 \Rightarrow t(\mathbf{A})t(v) = \lambda t(v) \Rightarrow \sigma(\mathbf{A}) \subseteq \sigma(t(\mathbf{A}))$. On the other hand $t(\mathbf{A})w = \alpha w \Rightarrow \mathbf{E}_m\mathbf{A}\mathbf{E}_m w =$

$\alpha w \Rightarrow \mathbf{E}_m\mathbf{E}_m\mathbf{A}\mathbf{E}_m w = \mathbf{E}_m \alpha w \Rightarrow \mathbf{A}\mathbf{E}_m w = \alpha \mathbf{E}_m w \Rightarrow \mathbf{A}\mathbf{E}_m w \mathbf{E}_1 = \alpha \mathbf{E}_m w \mathbf{E}_1 \Rightarrow \mathbf{A} t(w) = \alpha t(w)$ that means that $t(w)$ is an eigenvector for $\mathbf{A}$ corresponding to the eigenvalue α . Hence $\sigma(t(\mathbf{A})) \subseteq \sigma(t(\mathbf{A}))$ that together with $\sigma(\mathbf{A}) \subseteq \sigma(t(\mathbf{A}))$ complete the proof. Further results for centrosymmetric matrices can be found in Nield (1994).

Projections. It follows by **4** and **7** that if $\mathbf{P}$ is an orthogonal projection matrix, *i.e.* symmetric and idempotent, then $t(\mathbf{P})$ is an orthogonal projection matrix too. In particular, if $\mathbf{P} \in \mathbb{R}^{n\times n}$ is an orthogonal projection matrix on a subspace $\mathcal{S} \subseteq \mathbb{R}^n$, then $t(\mathbf{P}) \in \mathbb{R}^{n\times n}$ is the orthogonal projection matrix onto the subspace $t(\mathcal{S}) \subseteq \mathbb{R}^n$ spanned by a set of all the t-transformed vectors of a basis of $\mathcal{S}$. In fact, let $\mathcal{S} = span\{\mathbf{a}_1, \cdots, \mathbf{a}_p\} = \mathcal{C}(\mathbf{A})$, where $\mathbf{A} \in \mathbb{R}^{n\times p}, p < n$, is the matrix whose columns span $\mathcal{S}$. Hence $\mathbf{P} = \mathbf{A}\left(\mathbf{A}^T\mathbf{A}\right)^{-1}\mathbf{A}^T$ and by **4**, **6** and **9**, $t(\mathbf{P}) = t\left(\mathbf{A}\left(\mathbf{A}^T\mathbf{A}\right)^{-1}\mathbf{A}^T\right) = t(\mathbf{A})\left[t\left(\mathbf{A}^T\mathbf{A}\right)\right]^{-1} t(\mathbf{A})^T =$

$= t(\mathbf{A})\left[t(\mathbf{A})^T t(\mathbf{A})\right]^{-1} t(\mathbf{A})^T$ that is an orthogonal projection matrix onto the subspace

$t(\mathcal{S}) = \mathcal{C}(t(\mathbf{A})) = span\{t(\mathbf{a}_1), \cdots, t(\mathbf{a}_p)\}$.

Convolutions. The important properties in time series filtering are **6** and **7**, since matrix product is equivalent to linear filters convolution. As we will see in the following, the t-transformation allows an easy and useful description of the structure of the smoothing matrix that represents the transformation acting on the data to produce smooth estimates of the latent variables. The relevance of property **6** is that if $\mathbf{A}$ and $\mathbf{B}$ are smoothing matrices of a given structure, than $\mathbf{AB}$ is still a smoothing matrix and it conserves the structure of $\mathbf{A}$ and $\mathbf{B}$. The same holds when repeatedly smoothing a vector of observations by the same filter; in this case, property **7** is applied. These properties are crucial for the construction and study of filters resulting from the convolution of well-known systems of

weights.

3 Smoothing matrices and the role of the t-transformation

In time series analysis, a useful way to estimate the trend underlying the data is by fitting *locally* a polynomial function, such that any fitted value at a time point t depends only on the observations corresponding to time points in some specified neighborhood of t. Such a fitting curve is smooth by construction. Let us denote a time series as the set

$$\{(t_j, y_j), j = 1, \cdots, N\}$$

where each target point t_j is the time the observation y_j is taken. Any transformation s acting on the time series to produce smooth estimates is a *smoother*.

Usually s depends on a *smoothing parameter*, say η, which is selected according to the variability of the data and the amount of smoothing desired. The value of the smoothing parameter determines the number of observations averaged to obtain each estimate. In particular, if $\eta \to 0$, then the neighborhoods are made of only one observation and the result of the smoothing is an interpolation, whereas if $\eta \to \infty$, then all the observations are considered and smoothing produces a constant line corresponding to the mean of the series.

Any *linear* smoother can be represented by a squared matrix, let us call it $\mathbf{S}$, in such a way that

$$\begin{aligned} s &: \mathbb{R}^N \to \mathbb{R}^N \\ \mathbf{y} &\longmapsto \widehat{\mathbf{y}} = \mathbf{S}\mathbf{y} \end{aligned}$$

where $\mathbf{y} \in \mathbb{R}^N$ is an N-dimensional vector corresponding to the input data and $\widehat{\mathbf{y}} \in \mathbb{R}^N$ is the N-dimensional vector representing the smoothed values.

As long as s depends on the smoothing parameter, so does $\mathbf{S}$ and the relation between the original series and the corresponding smoothed estimates becomes

$$\widehat{\mathbf{y}} = \mathbf{S}_\eta \mathbf{y}.$$

It is crucial to remark that linearity holds only for fixed η. In fact, smoothers such that their smoothing parameters are selected by means of data dependent optimization criteria are non-linear.

Let now $w_{hj}, h, j = 1, \cdots, N$, denote the generic element of the *smoothing matrix* $\mathbf{S}_\eta$. The w_{hj}'s are the weights to be applied to the observations $y_j, j = 1, ..., N$, to get the estimate $\widehat{y}_h$, for each $h = 1, ..., N$, *i.e.*

$$\widehat{y}_h = \sum_{j=1}^{N} w_{hj} y_j.$$

These weights depend on the shape of the weight function associated to any smoother. Once the smoothing parameter has been selected, the w_{hj}'s for the observations corresponding to points falling out of the neighborhood of any target point are null, such that the estimates of the $N-2m$ central observations are obtained by applying $2m+1$ symmetric weights to the observations neighboring the target point. The estimates of the first and last m observations can be obtained by applying asymmetric weights of variable length to the first and last m observations

respectively, *i.e.*

$$\widehat{y}_h = \sum_{j=-m}^{m} w_{h,h-j} y_{h-j}, \quad h = m+1, \ldots, N-m \tag{3}$$
(central observations)

$$\widehat{y}_p = \sum_{r=1}^{m_p} w_{pr} y_r, \quad p = 1, \ldots, m \tag{4}$$
(initial observations)

$$\widehat{y}_q = \sum_{z=1}^{m_q} w_{q,N+1-z} y_{N+1-z}, \quad q = N-m+1, \ldots, N \tag{5}$$
(final observations)

where $2m+1$ is the length of the time invariant symmetric filter and m_p and m_q are the time-varying lengths of the asymmetric filters.

Hence, the smoothing matrix $\mathbf{S}_\eta$ has the following structure,

$$\mathbf{S}_\eta = \begin{bmatrix} \underset{(m\times 2m)}{\mathbf{W}^a} & & \underset{(m\times N-2m)}{\mathbf{O}} \\ & \underset{(N-2m\times N)}{\mathbf{W}^s} & \\ \underset{(m\times N-2m)}{\mathbf{O}} & & \underset{(m\times 2m)}{\mathbf{W}^{a'}} \end{bmatrix} \tag{6}$$

where $\mathbf{O}$ is a null matrix and $\mathbf{W}^a$, $\mathbf{W}^s$, $\mathbf{W}^{a'}$ are submatrices whose dimensions are shown in parentheses.

In particular, $\mathbf{W}^s$ is a $(2m+1)$-diagonal matrix (in the same sense of a tridiagonal matrix) and its row elements are the *symmetric* weights $w_{h,h-j}$ in (3) while the rows of the matrices $\mathbf{W}^a$ and $\mathbf{W}^{a'}$ are the sets of *asymmetric* weights w_{pr} and $w_{q,N+1-z}$ in (4) and (5) respectively. The length of the symmetric filter, as well as that of the asymmetric filters, depends on: (a) the shape of the smoother chosen to fit the data, (b) the value of the smoothing parameter and, (c) the number of decimals chosen for each weight, if the smoother is a continuous function, *e.g.*, as we will see, the Gaussian kernel.

In the following section we highlight the role of the t-transformation in the analysis of the smoothing matrices.

4 The t-transformation in time series filtering

Smoothing matrices of the form (6) are centrosymmetric, *i.e.*

$$t\left(\mathbf{S}_\eta\right) = \mathbf{S}_\eta \tag{7}$$

and their submatrices of symmetric weights are rectangular centrosymmetric,

$$\mathbf{W}^s = t\left(\mathbf{W}^s\right). \tag{8}$$

Furthermore, the submatrices of asymmetric weights for the first and last observations are the t-transform of each other,

$$\mathbf{W}^a = t\left(\mathbf{W}^{a'}\right). \tag{9}$$

The consequences of the above relations are important from both computational and theoretical viewpoints. In fact, relation (9) allows to halve the dimension of any smoothing problem by considering only m instead of $2m$ asymmetric filters. In particular, this reduction is substantial, especially when dealing with long filters that asymmetrically weight a considerable number of initial and end observations. On the other hand, theoretically, it becomes significant when asymmetric weights are derived on the

basis of assumptions that are different from those corresponding to symmetric weights. It is desirable to have asymmetric filters for end points that converge monotonically to the symmetric ones.

Examples of well-known smoothers for fixed values of the smoothing parameter of the form (6) are: (1) the locally weighted regression, or loess, that uses local polynomials of a degree d generally estimated by ordinary or weighted least squares and, hence, satisfies the criterium of best fit to the data; (2) the cubic smoothing splines, which search for an optimal compromise between the degree of fitting to the data and that of smoothing with respect to a second degree polynomial; (3) the Gaussian kernel, a locally weighted average where the weight function is the Gaussian standard distribution; and (4) the Henderson filters, whose weights derive from the graduation theory and minimize smoothing with respect to a third degree polynomial within the span of the filter. A detailed mathematical description of each smoother and the derivation of its weights are given in Dagum and Luati [5] and [6].

Next, we briefly introduce the centrosymmetric smoothing matrices associated to each of the above smoothers.

The symmetric and asymmetric weights of loess are given by the $1 \times (2m+1)$ row vector (Dagum and Luati [5])

$$\mathbf{S}_j = \mathbf{t}_j^T \left(\mathbf{T}_j^T \mathbf{W}_j \mathbf{T}_j\right)^{-1} \mathbf{T}_j^T \mathbf{W}_j$$

where $\mathbf{t}_j$ is a $(d+1)$-dimensional row vector of elements of the target point t_j with exponents $0, 1, \ldots, d$; $\mathbf{T}_j$ is an $(2m+1) \times (d+1)$ matrix of points belonging to the neighborhood of t_j with exponents $0, 1, \ldots, d$; and $\mathbf{W}_j$ is an $(2m+1) \times (2m+1)$ matrix of the weights for the observations corresponding to the points in the neighborhood of t_j.

The influential matrix (Wahba [22]) associated to a fixed smoothing parameter λ_0 of a cubic smoothing spline is given by

$$S_{\lambda_0} = \left[\mathbf{I}_N - \mathbf{D}^T \left(\frac{1}{\lambda_0} \mathbf{B} + \mathbf{D}\mathbf{D}^T \right)^{-1} \mathbf{D} \right]$$

where $\mathbf{I}_N$ is the $N \times N$ identity matrix and $\mathbf{B}$ and $\mathbf{D}$ are $(N-2) \times (N-2)$ and $(N-2) \times N$ matrices, respectively (see Dagum and Capitanio [4]).

For the Gaussian kernel (see Wand and Jones [21]), the generic weights w_{hj} to be applied to the observations y_j, $j = 1, \ldots, N$, to get the estimate of the observation y_h, are given by

$$w_{hj} = \frac{\frac{1}{\sqrt{2\pi}} \exp\left\{ -\frac{1}{2} \left(\frac{t_h^* - t_j}{b} \right)^2 \right\}}{\sum_{i=1}^{N} \frac{1}{\sqrt{2\pi}} \exp\left\{ -\frac{1}{2} \left(\frac{t_h^* - t_i}{b} \right)^2 \right\}}$$

for $h = 1, \ldots, N$.

The symmetric weights of the 9-term Henderson filter can be obtained from the general formula (Henderson [8])

$$h_n = \frac{315\left[(q-1)^2 - n^2\right](q^2 - n^2)\left[(q+1)^2 - n^2\right](3q^2 - 16 - 11n^2)}{8q(q^2-1)(4q^2-1)(4q^2-9)(4q^2-25)}$$

by making $q = 6$ such that the values h_n are obtained for each n from -4 to 4. The explicit form of the asymmetric weights, which do not follow the above assumptions, is derived by Laniel [13] based on the work of Musgrave [16].

To illustrate the effect of the t-transformation in time series filtering, we consider the Gaussian kernel and the Henderson filter. Fig. 1 and 2 show the smoothing matrix (6) associated to a 13-term Gaussian kernel and to a 9-term Henderson filter

0.347	0.302	0.199	0.100	0.038	0.011	0.002	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
0.232	0.267	0.232	0.153	0.077	0.029	0.008	0.002	0.000	0.000	0.000	0.000	0.000	0.000	0.000
0.133	0.201	0.231	0.201	0.133	0.066	0.025	0.007	0.002	0.000	0.000	0.000	0.000	0.000	0.000
0.062	0.125	0.189	0.217	0.189	0.125	0.062	0.024	0.007	0.001	0.000	0.000	0.000	0.000	0.000
0.023	0.061	0.122	0.184	0.212	0.184	0.122	0.061	0.023	0.007	0.001	0.000	0.000	0.000	0.000
0.007	0.023	0.060	0.121	0.183	0.210	0.183	0.121	0.060	0.023	0.007	0.001	0.000	0.000	0.000
0.001	0.007	0.023	0.060	0.121	0.183	0.210	0.183	0.121	0.060	0.023	0.007	0.001	0.000	0.000
0.000	0.001	0.007	0.023	0.060	0.121	0.183	0.210	0.183	0.121	0.060	0.023	0.007	0.001	0.000
0.000	0.000	0.001	0.007	0.023	0.060	0.121	0.183	0.210	0.183	0.121	0.060	0.023	0.007	0.001
0.000	0.000	0.000	0.002	0.007	0.023	0.060	0.121	0.183	0.210	0.183	0.121	0.060	0.023	0.007
0.000	0.000	0.000	0.000	0.001	0.007	0.023	0.061	0.122	0.184	0.212	0.184	0.122	0.061	0.023
0.000	0.000	0.000	0.000	0.000	0.001	0.007	0.024	0.062	0.125	0.189	0.217	0.189	0.125	0.062
0.000	0.000	0.000	0.000	0.000	0.000	0.002	0.007	0.025	0.066	0.133	0.201	0.231	0.201	0.133
0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.002	0.008	0.029	0.077	0.153	0.232	0.267	0.232
0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.002	0.011	0.038	0.100	0.199	0.302	0.347

Figure 1: Smoothing matrix $\mathbf{S}$ of the 13-term Gaussian kernel smoother for a time series of 15 data points.

to be applied to series of 15 observations, with values approximated to three decimals. It is evident that these matrices are centrosymmetric as well as those with the symmetric weights. The submatrices with the asymmetric weights for first and last observations are the t-transform of each other. Fig. 3 and 4 illustrate the path of the asymmetric and symmetric weights of, respectively, the 13-term Gaussian kernel and the 9-term Henderson filter for the last observations. It is apparent that the asymmetric weights for the last observations monotonically converge to the symmetric weights.

0.581	0.424	0.185	-0.034	-0.156	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
0.298	0.354	0.282	0.126	-0.011	-0.049	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
0.086	0.242	0.315	0.259	0.120	0.000	-0.022	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
-0.029	0.102	0.255	0.324	0.263	0.120	-0.004	-0.031	0.000	0.000	0.000	0.000	0.000	0.000	0.000
-0.041	-0.010	0.119	0.267	0.330	0.267	0.119	-0.010	-0.041	0.000	0.000	0.000	0.000	0.000	0.000
0.000	-0.041	-0.010	0.119	0.267	0.330	0.267	0.119	-0.010	-0.041	0.000	0.000	0.000	0.000	0.000
0.000	0.000	-0.041	-0.010	0.119	0.267	0.330	0.267	0.119	-0.010	-0.041	0.000	0.000	0.000	0.000
0.000	0.000	0.000	-0.041	-0.010	0.119	0.267	0.330	0.267	0.119	-0.010	-0.041	0.000	0.000	0.000
0.000	0.000	0.000	0.000	-0.041	-0.010	0.119	0.267	0.330	0.267	0.119	-0.010	-0.041	0.000	0.000
0.000	0.000	0.000	0.000	0.000	-0.041	-0.010	0.119	0.267	0.330	0.267	0.119	-0.010	-0.041	0.000
0.000	0.000	0.000	0.000	0.000	0.000	-0.041	-0.010	0.119	0.267	0.330	0.267	0.119	-0.010	-0.041
0.000	0.000	0.000	0.000	0.000	0.000	0.000	-0.031	-0.004	0.120	0.263	0.324	0.255	0.102	-0.029
0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	-0.022	0.000	0.120	0.259	0.315	0.242	0.086
0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	-0.049	-0.011	0.126	0.282	0.354	0.298
0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	-0.156	-0.034	0.185	0.424	0.581

Figure 2: Smoothing matrix **S** of the 9-term Henderson filter for a time series of 15 data points.

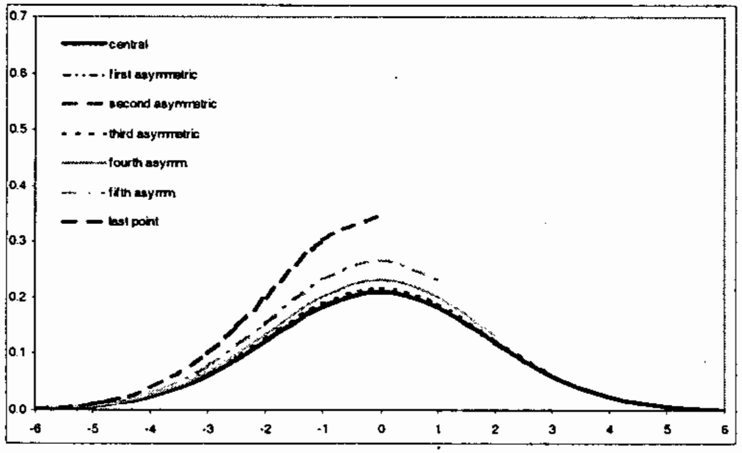

Figure 3: Convergence of asymmetric to the symmetric weights of the 13-term Gaussian kernel smoother.

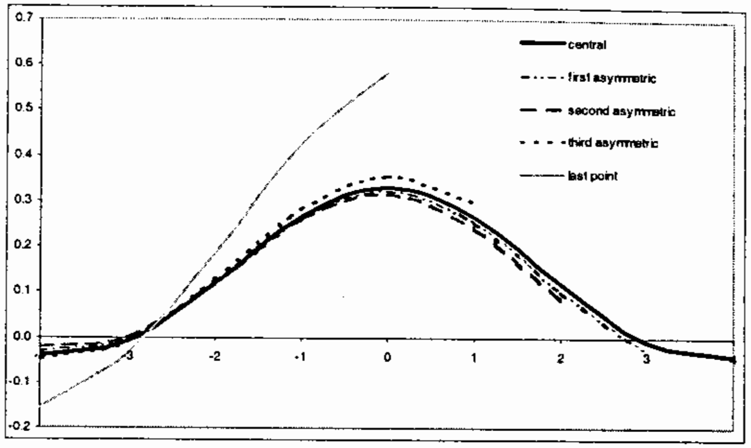

Figure 4: Convergence of asymmetric to the symmetric weights of the 9-term Henderson filter.

References

[1] Andrew, A.L. 1998. Centrosymmetric matrices. *SIAM Rev.* 40, No. 3: 697-698.

[2] Collar, A. R. 1962. On Centrosymmetric and Centroskew Matrices. *Quart. J. Mech. Appl. Math.* 15:265-281.

[3] Chen, H. -C. 1998. Generalized Reflexive Matrices: Special Properties and Applications. *SIAM. J. Matrix. Anal. Appl.* 19, No 1: 140-153.

[4] Dagum, E. B. and Capitanio, A. 1999. Cubic Splines Spectral Properties for Short Term Trend-Cycle Estimation. *Proceedings of the Business and Economic Statistics Section of the American Statistical Association* 100-105.

[5] Dagum, E.B. and Luati, A. 2000. Predictive Performances of Some Linear and Nonlinear Smoothers for Noisy Data. *Statistica*, anno LX, 4:635-654.

[6] Dagum, E. B. and Luati, A. 2001. A Study of the Asymmetric and Symmetric Weights of Kernel Smoothers and their Spectral Properties, *Estadistica* (forthcoming).

[7] Farebrother, R. W. 2000. Note on the Elementary Properties of Reflection and Permutation Matrices, in *Multivariate Statistical Analysis, A Festschrift for Heinz Neudecker* (Heijmans, R.D.H., Pollock D.S.G. and Satorra, A., Eds.), Kluwer Academic Publisher.

[8] Henderson, R. 1916. Note on Graduation by Adjusted Average. *Transaction of the Actuarial Society of America.* 17:43-48.

[9] Kimura, M. 1957. Some Problems of Stochastic Processes in Genetics. *Annals of Mathematical Statistics*, 28:82-901.

[10] Kontoghiorghes, E. J. 2000. *Parallel Algorithms for Linear Models: Numerical Methods and Estimation Problems.* Advances in Computational Economics, 15, Kluwer Academic Publishers.

[11] Krafft, O. and Schaever, M. 2000. Centrortogonal Matrices. *Linear Algebra Appl.* 306:145-154.

[12] Iosifescu, M. 1980. *Finite Markov Processes and their Applications*, Wiley.

[13] Laniel, N. 1986. Design Criteria for the 13-term Henderson End Weights. *Working Paper TSRA 1986-011*, Statistics Canada.

[14] Magnus, J.R. and Neudecker, H. 1979. The Commutation Matrix: Some Properties and Applications. *The Annals of Statistics*, 7:381-394.

[15] Melman, A. 1999. Centrosymmetric Matrix-Vector Multiplication. *Technical Report MATH-1999-02, March*, Department of Mathematics, University of San Francisco.

[16] Musgrave, J.C. 1964. A Set of End Weights to End All End Weights. *Unpublished Internal Note*, Washington, US Bureau of Census.

[17] Nield, D. A. 1994. Odd-even Factorization Results for Eigenvalue Problems. *SIAM Review* 36:649-651.

[18] Trench, W. F. 1990. Spectra Evolution of a One-parameter Extension of a Real Symmetric Toeplitz Matrix. *SIAM, J. Matrix Anal. Appl.* 11:601-611.

[19] Trench, W. F. 1993. Interlacement of the Even and Odd Spectra of Real Symmetric Toeplitz Matrices. *Linear Algebra. Appl.* 195:59-68.

[20] Trench, W. F. 1997. Numerical Solution of the Inverse Eigenvalue Problem for Real Symmetric Toeplitz Matrices. *SIAM, J. Sci. Comput.*18:1722-1736.

[21] Wand, M.P. and Jones, M.C. 1995. *Kernel Smoothing.* Monographs on Statistics and Applied Probability, 60, Chapman&Hall.

[22] Wahba, G. 1990. *Spline Models for Observational Data.* SIAM Philadelphia.

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