JOURNAL OF APPROXIMATION THEORY 33, 43–58 (1981)

Moving-Weighted-Average Smoothing Extended to the Extremities of the Data
III. Stability and Optimal Properties*

T. N. E. GREVILLE

Mathematics Research Center, University of Wisconsin, Madison, Wisconsin 53706

Communicated by Oved Shisha

Received April 21, 1980

1. INTRODUCTION

A time-honored method of smoothing equally spaced observations (such as time series, or human mortality rates by age) to remove or reduce unwanted irregularities is the moving weighted average (MWA). Here we shall consider only symmetrical averages, which can be expressed in the form [15]

$$u_x = \sum_{j=-m}^{m} c_j y_{x-j} \qquad (c_{-j} = c_j). \tag{1.1}$$

where y_x is the observed value corresponding to the (integral) argument x, u_x is the corresponding smoothed value, m is a given positive integer, and the coefficients c_j are given real quantities satisfying

$$\sum_{j=-m}^{m} c_j = 1.$$

This smoothing method has the disadvantage that it does not produce smoothed values for the first m and the last m observations, unless the original data set can be extended by m observations at each end. In the first paper of this series [9], I suggested a natural method of extending the smoothing to the extremities of the data, in which the treatment of the observations in the "tails" of the data set is an integral part of an overall matrix-vector operation and not something extra grafted on at the ends. In fact,

$$u = Gy, \tag{1.2}$$

where y is the vector of observed values, u is the corresponding vector of

* Sponsored by the United States Army under Contract No. DAAG-29-75-C-0024.

0021-9045/81/090043-16$02.00/0

smoothed values, and G is a square matrix uniquely determined by the following five properties.

(1) The components of u, except the first m and the last m, are merely the smoothed values that would be obtained by application of the given MWA.

(2) $G = (g_{ij})$ of order N is a band matrix such that $g_{ij} = 0$ for $|i-j| > m$, and $N > 2m$.

(3) Let (1.1) be exact for polynomials of degree $2s-1$, but not, in general, for those of higher degree. (Note that, because of the symmetry of (1.1), the degree of exactness must be odd.) Also let K denote the matrix of $N-s$ rows and N columns that transforms a vector u into the vector Ku of sth finite differences of the components of u. Then, G is of the form

$$G = I - K^T DK \tag{1.3}$$

for some D of order $N-s$.

(4) D is nonsingular and has a Toeplitz inverse.

(5) If $D^{-1} = (t_{ij})$, with $t_{ij} = t_{j-i}$, then the series

$$\sum_{\nu=-\infty}^{\infty} t_\nu z^\nu$$

converges in some part of the complex plane. (Note that it was shown in [11] that t_ν depends only on ν and is independent of N.)

The existence and uniqueness of G [9, Theorem 3.1] depend on certain (very mild) hypotheses concerning the given MWA (1.1). Following the customary notation of the calculus of finite differences, we define E and δ by

$$Ef(x) = f(x+1), \qquad \delta f(x) = f(x+\tfrac{1}{2}) - f(x-\tfrac{1}{2}). \tag{1.4}$$

Then (1.1) can be written in the form

$$u_x = [1 - (-1)^s \delta^{2s} q(E)]\, y_x, \tag{1.5}$$

where

$$q(E) = \sum_{j=-m+s}^{m-s} q_j E^j, \tag{1.6}$$

with $q_{-j} = q_j$, and we impose the two conditions that $q_0 > 0$ and that $q(z)$ have no zero on the unit circle of the complex plane. The effect of these hypotheses is discussed in [9], where the matrix-vector formulation is also shown to be equivalent to a certain extrapolation algorithm. In [10] efficient numerical procedures are described.

It is the purpose of this paper to show that if the given MWA is further restricted, the unique G determined by the five properties listed has a certain stability property and certain optimal properties.

2. Trench Matrices

A matrix $L = (\ell_{ij})_{i,j=0}^{N}$ will be called *strictly banded* if $\ell_{ij} = 0$ for $j - i > h$ and for $i - j > k$, where h and k are nonnegative integers and $h + k \leqslant N$. Note that the numbering of rows and columns is started with 0 rather than 1. Let

$$L_i(x) = \sum_{j=0}^{N} \ell_{ij} x^j$$

be the generating function of the elements of the ith row of L. In [8], I defined a Trench matrix as a strictly banded matrix such that

$$\begin{aligned} L_i(x) &= x^i A(x) \sum_{\mu=0}^{i} b_\mu x^{-\mu} && (0 \leqslant i < k) \\ &= x^i A(x)\, B(1/x) && (k \leqslant i \leqslant N - h) \\ &= x^i B(1/x) \sum_{\nu=0}^{N-i} x^\nu && (N - h < i \leqslant N), \end{aligned} \tag{2.1}$$

where

$$A(x) = \sum_{\nu=0}^{h} a_\nu x^\nu, \qquad B(x) = \sum_{\mu=0}^{k} b_\mu x^\mu \tag{2.2}$$

are polynomials with real or complex coefficients (according as L is real or complex) and $a_0 b_0 \neq 0$.

This form was previously given (though the name "Trench matrix" was not used) in the joint paper [11], and the properties of such matrices were studied. They were studied further in [7, 8], and in [9] the results were applied to smoothing matrices G of the form (1.3). In [11] it was shown that a strictly banded matrix L has a Toeplitz inverse if and only if it is a nonsingular Trench matrix, and further that the Trench matrix (2.1) is nonsingular if and only if $A(x)$ and $B(x)$ have no common zero. For convenience the following results based on [7, 9] are stated as lemmas.

Lemma 2.1. *Let a given symmetrical* MWA *of* $2m + 1$ *terms be exact for the degree* $2s - 1$ $(s \leqslant m)$, *let* $q(x)$ *defined by* (1.5) *and* (1.6) *have no* *zeros on the unit circle, and let* $q_0 > 0$. *Then there exist at least one and at most a finite number not exceeding* 2^{m-s} *of polynomials* $A(x)$ *of degree* $m-s$ *with real coefficients such that*

$$q(x) = A(x)\,A(1/x). \tag{2.3}$$

LEMMA 2.2. *Let a given symmetrical* MWA *satisfy the hypotheses of Lemma* 2.1 *and let* G *be real and symmetric of order at least* $2m+1$ *and have properties* (1)–(4) *of Section* 1. *Then* $G = I - F$, *where* F *is the symmetric Trench matrix characterized by two identical polynomials* $\hat{A}(x)$ *and* $\hat{B}(x)$ *both equal to* $(x-1)^s A(x)$, *where* $A(x)$ *has real coefficients and satisfies* (2.3).

Of the five properties utilized in Section 1 to define the matrix G uniquely, property (1) is no more than a restatement of the problem to be solved, while property (2) requires, in effect, that the overall smoothing procedure be a "local" one. Properties (3)–(5) were motivated in [9] by the fact they are also properties of the well-known Whittaker smoothing method (for a description of the latter see [9, 5]). Of these property (5) may appear to some readers more artificial and less compelling than the others. It is therefore of interest to know that it can be replaced by certain alternative conditions. Two of these are described in the following two sections.

It follows from Lemmas 2.1 and 2.2 that, in general, properties (1)–(4) of Section 1 do not uniquely determine G but restrict it to a finite class. The various matrices of the class differ, however, only in the square submatrices of order m in the upper left and lower right corners. Moreover, we shall find it convenient, in Theorems 3.1 and 5.1, to introduce the additional hypothesis that G be symmetric. The reader will note that the five properties of Section 1 do not explicitly mention symmetry of G; rather this comes out as a consequence, and property (5) plays an essential role (see proof of Theorem 3.1 of [9]). I conjecture that the symmetry of G is not a necessary hypothesis, but I have not found satisfactory proofs without it. It will be noted that, in consequence of the first four properties, symmetry obtains everywhere except possibly in the two corner submatrices referred to (see [9]).

It may be mentioned, in fact, that there are cogent reasons for thinking that G should be symmetric. A square matrix is called *persymmetric* if it is symmetric about its secondary diagonal. It is called *centrosymmetric* if it is symmetric about the center of the matrix: thus $G = (c_{ij})$ of order N is centrosymmetric if $c_{ij} = c_{N-j+1,N-i+1}$ for all (i,j). Now, it is easily seen that of the three properties of symmetry, persymmetry, and centrosymmetry, any two imply the third. If G has properties (1)–(4), it is necessarily persymmetric, because $G = I - F$, where F is a Trench matrix, and every Trench matrix is persymmetric [11]. Therefore, if G is not symmetric, it is not centrosymmetric. Now, if G is not centrosymmetric, this means that reversing the order of the observed values (without changing their magnitude) would not merely reverse the order of the smoothed values, but would cause different numerical values to be obtained. For example, the elements of the bottom row row of G would not be those of the top row in reverse order. The formula for smoothing the last observation would not be the mirror image of the one for smoothing the first observation, but would be a different formula. This would seem to be an undesirable characteristic of the smoothing process.

3. Optimal Property of the Extreme Rows of G

The atypical rows of G (i.e., the first m and the last m) may be regarded as representing special unsymmetrical smoothing formulas employed near the ends of the data to supplement the main symmetrical formula used elsewhere. Under certain statistical hypotheses concerning the observations subjected to smoothing (see, e.g., [1, 10]), the Euclidean length of each row of G may be regarded as the ratio of reduction in the standard deviation of error of the particular observation that results from the smoothing operation. It would therefore be desirable to choose from the finite class of possible matrices G one for which the Euclidean length of the atypical rows is small. In particular, Tables 4–6 of [10] suggest that there is a strong tendency for the length of the top and bottom rows to become large. Because the bottom row contains the same elements as the top row but in reverse order, it is sufficient to consider the top row.

Theorem 3.1. *Let a given symmetrical* MWA *satisfy the hypotheses of Lemma* 2.1 *and also* $c_0 > -1$, *where* c_0 *is defined by* (1.1). *Let* G *be symmetric and have properties* (1)–(4) *of Section* 1. *Then the Euclidean length of the top row of* G *is smallest when* G *has property* (5).

Proof. Let G be symmetric and have properties (1)–(4). Since $q_0 > 0$, it follows from Lemma 2.2 that $G = I - F$, where F is the (singular) Trench matrix characterized by the two identical polynomials

$$\hat{A}(x) = \hat{B}(x) = (x-1)^s A(x),$$

and $A(x)$ satisfies (2.3). Because of the symmetry of the coefficients q_j of (1.6) and the fact that $q(x)$ has no zeros on the unit circle, the polynomial $x^{m-s}q(x)$ of degree $2m-2s$ has $m-s$ zeros inside the unit circle and an equal number outside, which are the reciprocals of those inside. The different possible choices of $A(x)$ correspond to the different ways of assigning the $2m-2s$ zeros to two equal subsets in such a way that a given zero and its reciprocal are not in the same subset. Conjugate pairs of complex zeros must be assigned to the same subset.

Let

$$\hat{A}(x) = \sum_{j=0}^{m} \hat{a}_j x^j.$$

Then, (1.5) can be written in the form

$$u_x = [1 - \hat{A}(E)\hat{A}(E^{-1})] y_x, \tag{3.1}$$

and so, by (1.1), $c_0 = 1 - S$, where

$$S = \sum_{j=0}^{m} \hat{a}_j^2.$$

Therefore,

$$S = 1 - c_0. \tag{3.2}$$

By (2.1), the nonzero elements of the top row of G are $1 - \hat{a}_0^2$, $-\hat{a}_0\hat{a}_1$, $-\hat{a}_0\hat{a}_2, \ldots, -\hat{a}_0\hat{a}_m$. If R_0 denotes the Euclidean length, we have

$$R_0^2 = 1 - 2\hat{a}_0^2 + \hat{a}_0^2 S = 1 - \hat{a}_0^2(2 - S). \tag{3.3}$$

If $r_1, r_2, \ldots, r_{m-s}$ are the zeros of $A(x)$,

$$\hat{A}(x) = \hat{a}_m (x-1)^s \prod_{j=1}^{m-s} (x - r_j). \tag{3.4}$$

If

$$P = (-1)^{m-s} \prod_{j=1}^{m-s} r_j, \tag{3.5}$$

we have, by (3.1) and (3.4),

$$\hat{a}_0 = -\hat{a}_m P, \qquad c_m = -\hat{a}_0 \hat{a}_m,$$

and therefore,

$$\hat{a}_0^2 = c_m P. \tag{3.6}$$

In view of (3.2) and (3.6), (3.3) becomes

$$R_0^2 = 1 - c_m P(1 + c_0).$$

In this relation c_0 and c_m are given; P is the only variable. Moreover, $c_m P = \hat{a}_0^2$ is positive, and $1 + c_0$ is positive since $c_0 > -1$ by hypothesis. Therefore R_0^2 is smallest when $|P|$ is largest, and by (3.5) this is clearly the case when $A(x)$ is chosen so that its $m - s$ zeros are the zeros of $q(x)$ that are outside the unit circle. But by Theorem 3.1 of [9], this choice of $A(x)$ is equivalent to property (5). This completes the proof.

I conjecture that the hypothesis that G be symmetric is not essential. However, dropping it might necessitate stronger conditions on the MWA.

4. Characteristic Function of an MWA

Schoenberg [15] defined the characteristic function of the MWA (1.1) as

$$\phi(t) = \sum_{j=-m}^{m} c_j e^{ijt}. \tag{4.1}$$

For a symmetrical MWA this is a real function of the real variable t, and can be expressed in the alternative form

$$\phi(t) = \sum_{j=-m}^{m} c_j \cos jt.$$

It is periodic with period 2π and equal to unity for $t = 2\pi n$ for all integers n.

The effect of MWA's in eliminating or reducing certain waves is well known (e.g., [4, 12]). If the input to the smoothing process is a sine wave, which can be represented in the form

$$y_x = C\cos(rx + h), \tag{4.2}$$

it can be shown by simple algebraic manipulation that

$$u_x = y_x \phi(2\pi/\omega),$$

where $\omega = 2\pi/r$ is the period of y_x. Thus, if $\phi(2\pi/\omega) = 0$, the wave is annihilated by the smoothing process; the amplitude is severely reduced if it is close to zero. Thus MWA smoothing is related to the "filtering" processes considered by Wiener [19] and others.

Schoenberg [15] defined a *smoothing formula* as an MWA whose characteritsic function $\phi(t)$ satisfies the condition

$$|\phi(t)| \leqslant 1 \tag{4.3}$$

for all t. Thomeé [18] calls (4.3) "von Neumann's condition" without, however, citing any specific publication of von Neumann. Later Schoenberg [16, 17] suggested the stronger condition

$$|\phi(t)| < 1 \qquad (0 < t < 2\pi). \tag{4.4}$$

Lanczos (see [17]) pointed out that (4.4) is obtained by requiring that every simple vibration (4.2) be diminished in amplitude by the transformation (1.1).

The main theorem (Theorem 3.1) of [9] includes the hypothesis that the given MWA is such that $q_0 > 0$. The following theorem shows that this inequality follows from (4.3).

THEOREM 4.1. *Let a given symmetrical* MWA *exact for the degree* $2s-1$ *be such that* $q(x)$ *defined by* (1.5) *has no zeros on the unit circle, and let its characteristic function satisfy* (4.3). *Then* $q_0 > 0$.

Proof. Consider the real function $\psi(t) = 1 - \phi(t)$ and note that (4.3) is equivalent to

$$0 \leqslant \psi(t) \leqslant 2 \tag{4.5}$$

for all t. From (1.4), (1.5), and (4.1) it follows that

$$\psi(t) = (-1)^s \, (2i \sin \tfrac{1}{2}t)^{2s} \, q(e^{it}) = (4 \sin^2 \tfrac{1}{2}t)^s \, q(e^{it}),$$

and therefore (4.5) implies that $q(e^{it})$ is nonnegative for $0 < t < 2\pi$. In fact, it is positive, since $q(x)$ has no zeros on the unit circle, and by continuity it is positive for $t = 0$ as well. In other words, $q(1) > 0$. Now let the polynomials $A(x)$ and $B(x)$ be chosen so that $q(x) = A(x)\, B(1/x)$ and the zeros of $B(1/x)$ are the reciprocals of those of $A(x)$. This is always possible because of the symmetry of the coefficients of $q(x)$. Moreover, the coefficients in these polynomials can be normalized, as in the proof of Theorem 3.1 of [9], so that

$$q(x) = \pm A(x)\, A(1/x). \tag{4.6}$$

Then (4.6) and (2.2) give

$$q_0 = \pm \sum_{j=0}^{m-s} a_j^2, \tag{4.7}$$

and moreover,

$$q(1) = \pm |A(1)|^2.$$

Since we have shown that $q(1)$ is positive, the positive sign holds throughout, and (4.7) gives $q_0 > 0$, as required, since $a_0 \neq 0$.

5. The Stability Theorem

The matrix-vector approach to smoothing described in [9] and summarized above suggests an alternative to Schoenberg's criterion (4.3) or (4.4) for a smoothing formula. If the matrix G of (1.2) is defined in some unique way for all orders N greater than a certain minimum, and is therefore denoted by G_N, we call it *stable* if the limit

$$G_N^{\infty} = \lim_{n \to \infty} G_N^n$$

exists for all N. Schoenberg [17, Footnote 3] suggested a relationship between (4.3) and the conditions for the existence of the infinite power of a matrix [14, 3], but he did not elaborate the connection. In the theorem which follows we attempt to do so.

We summarize briefly some results of [8] that will be needed in the proof. If the polynomials that characterize a real symmetric Trench matrix H are $A(x)$ and $B(x)$ of degree d, the coefficients can be normalized so that either $B(x) = A(x)$ or $B(x) = -A(x)$. If the minus sign holds, one can consider the symmetric Trench matrix $-H$. It is sufficient, therefore, to consider the case in which $B(x) = A(x)$.

Let $A(x)$ be given and consider the family of symmetric matrices H_N of order $N \geqslant 2d + 1$ characterized by $A(x)$ and $B(x) = A(x)$. Let

$$G_N = I - \mu H_N,$$

where μ is a positive constant, and let

$$h(x) = A(x)\,A(1/x).$$

Then it is shown that $h(x)$ is real and nonnegative on the unit circle, and has a maximum thereon, which we denote by M, while Corollary 1 of [8] states that the family $\{G_N\}$ is stable if and only if

$$\mu \leqslant 2/M$$

and no zero of $A(x)$ is inside the unit circle unless it is also a zero of $A(1/x)$. A particular application of Lemma 1 of [8] yields the result that if D is a Trench matrix characterized by the polynomials $A(x)$ and $B(x)$, then $K^T DK$ (with K defined as in property (3) of Section 1 of the paper) is a (singular) Trench matrix characterized by the polynomials $\hat{A}(x) = (x-1)^s A(x)$ and $\hat{B}(x) = (x-1)^s B(x)$.

Theorem 5.1. *Let a symmetrical* MWA (1.1) *be given and let the associated smoothing matrix G_N for all $N \geqslant 2m + 1$ be symmetric and have* *properties* (1)–(4) *of Section* 1. *Then the family* $\{G_N\}$ *is stable if and only if* (4.3) *holds and the polynomial* $A(x)$ *associated with the matrix* D *of Property* (3) *has no zero inside the unit circle.*

Proof. From the hypotheses stated in the first sentence of the theorem we can deduce certain properties of the matrices $F = I - G$ and D. First we note that the hypotheses of the present theorem differ slightly from those of Theorem 3.1 of [9]. We have added the hypothesis that G is symmetric, and have omitted the restrictions on $q(x)$. However, the reader should note carefully that the latter omission is occasioned only by the fact that these restrictions are implied by the symmetry of G in conjunction with other hypotheses. The symmetry of G implies that of F. As the rows of K are linearly independent, it has full row rank and therefore has a left inverse, say J [2, Lemma 1.2]. Therefore,

$$J^t FJ = J^t K^t DKJ = D,$$

and consequently D is symmetric.

By property (2), D is strictly banded, and it follows from property (4) (see [11]) that D is a nonsingular Trench matrix. If it is characterized by the two polynomials $A(x)$ and $B(x)$ of degree $m - s$, then

$$q(x) = A(x)\, B(1/x),$$

as in the uniqueness proof of Theorem 3.1 of [9]. As D is real and symmetric, the coefficients in these polynomials are real and can be normalized so that

$$B(x) = \pm A(x). \tag{5.1}$$

As we have omitted the hypothesis that $q_0 > 0$, some ambiguity remains for the time being about the sign of the right member of (5.1), and we have

$$q(x) = \pm A(x)\, A(1/x). \tag{5.2}$$

Now, the symmetry and nonsingularity of D and the requirement that $A(x)$ have real coefficients imply that $q(x)$ has no zeros on the unit circle. As we have seen, symmetry of D implies (5.2), and nonsingularity implies [11] that $A(x)$ and $A(1/x)$ have no common zero. Now, if $q(x)$ has a zero on the unit circle, say x_0, then x_0^{-1} is also on the unit circle, so that $A(x)$ must have a zero on the unit circle; call it ρ. Then ρ^{-1} is a zero of $A(1/x)$, and $\bar{\rho}$ is a zero of $A(x)$, since $A(x)$ has real coefficients. But in this case $\bar{\rho} = \rho^{-1}$; therefore $A(x)$ and $A(1/x)$ have a common zero, a contradiction. Thus the supposition that $q(x)$ has a zero on the unit circle is false.

Now, suppose that (4.3) holds and $A(x)$ has no zeros inside the unit circle.

Then it follows from the proof of Theorem 4.1 that the positive sign holds in (4.6) and (4.7), and therefore in (5.2). By Lemma 1 of [8], F is a singular Trench matrix characterized by the polynomials

$$\hat{A}(x) = \hat{B}(x) = (x-1)^s A(x). \tag{5.3}$$

Let

$$f(x) = \hat{A}(x)\,\hat{B}(1/x) = (x^{1/2} - x^{-1/2})^{2s}\, q(x).$$

Then,

$$\psi(t) = f(e^{it}). \tag{5.4}$$

Let M denote the maximum value of $f(x)$ on the unit circle. Then by (4.5) $M \leqslant 2$, or

$$1 \leqslant 2/M. \tag{5.5}$$

Consequently, by Corollary 1 of [8], the family $\{G_N\}$ is stable.

Conversely, suppose that the family $\{G_N\}$ is stable, in addition to the hypotheses in the first sentence of the theorem. Since G_N is symmetric, its eigenvalues are real, and stability implies [14, 3] that all its eigenvalues are in the half-open interval $(-1, 1]$. In other words, all the eigenvalues of F_N are in $[0, 2)$ for all N. Now, if v is an arbitrary column vector of real elements, it is well known that the minimum value of the Rayleigh quotient $v^T Fv/v^T v$ is the (algebraically) smallest eigenvalue of F. Suppose the minus sign holds in (5.1) and let v be the unit vector with 1 as its first element and all the other elements 0. By (5.3) the constant term of $\hat{A}(x)$ is $(-1)^s a_0$, and the Rayleigh quotient is $-a_0^2$, which is negative since $a_0 \neq 0$ by the definition of a Trench matrix. Thus, F has a negative eigenvalue, in contradiction to the statement that all its eigenvalues are in $[0, 2)$. Therefore the supposition that the minus sign holds in (5.1) is false.

Since the positive sign holds in (5.1), F belongs to the class of matrices to which Corollary 1 of [8] applies. Thus stability of the family $\{G_N\}$ implies that $A(x)$ has no zero inside the unit circle unless it is also a zero of $A(1/x)$. But a common zero of $A(x)$ and $A(1/x)$ would imply that D is singular, which would contradict property (4). Therefore $A(x)$ has no zero inside the unit circle. Stability implies further that (5.5) holds, with M defined as before, and this implies in turn that $M \leqslant 2$, which, in view of (5.4), is tantamount to (4.5) and therefore to (4.3). This completes the proof.

It is easily verified that $G^{\mathscr{T}}$, when it exists, is the orthogonal projector on the eigenspace of G associated with the eigenvalue 1, that is, the space of N-vectors whose components are successive equally spaced ordinates of polynomials of degree $s-1$ or less.

I conjecture that in Theorem 5.1 the hypothesis of symmetry of G could be replaced by mild restrictions on the given MWA. While symmetry of the main part of G follows from the symmetry of the coefficients in the main formula and properties (1)–(4), the special corner submatrices are not symmetric unless $A(x)$ is is chosen so that $B(x) = A(x)$. I have not been able to find a stable family $\{G_N\}$ with unsymmetric corners, and I doubt that one exists. However, the proof of stability in [8] makes extensive use of the well-known relation between Rayleigh quotients and eigenvalues that holds only for Hermitian (including symmetric) matrices. Extension of the result to smoothing matrices with unsymmetric corners would require a different method of proof, which I have not succeeded in finding.

6. Smoothing Formulas in the Strict Sense and Another Optimal Property

Under certain conditions the smoothing procedure discussed here can be shown to minimize a certain "loss function" analogous to the Whittaker criterion (see [9]). In a slightly more general form of the Whittaker smoothing method [5] one minimizes the sum of the weighted squares of the departures of the smoothed values from the observed values plus a specified quadratic form in the sth differences of the smoothed values. In matrix terms this is

$$(u-y)^T W(u-y) + (Ku)^T HKu,$$

where W is a positive diagonal matrix and H is a given positive definite matrix of order $N-s$. This reduces to the usual Whittaker criterion when H is a scalar matrix gI. We shall consider here only the "unweighted" case $W = I$. Minimization of the "loss function" then leads to the equation

$$(I + K^T HK)\, u = y,$$

which has a unique solution for u since $I + K^T HK$ is positive definite. In [5] I showed that this smoothing method has the interesting property that if roughness (opposite of smoothness) is measured by the term $(Ku)^T HKu$, smoothness is always increased by the graduation. By Theorem 5.22 of [13],

$$(I + K^T HK)^{-1} = I - K^T(H^{-1} + KK^T)^{-1} K.$$

The last expression is of the form (1.3) and suggests that the use of an MWA with the natural extension of Section 1 might be regarded as a generalized Whittaker smoothing process if

$$D = (H^{-1} + KK^T)^{-1}.$$

Solving for H gives

$$H = (D^{-1} - KK^T)^{-1}. \tag{6.1}$$

We are led to inquire, therefore, under what conditions an MWA is such that the right member of (6.1) for the natural extension is positive definite for all N. We note in passing that

$$H^{-1} = D^{-1} - KK^T$$

is a Toeplitz matrix.

Schoenberg [15, p. 53] remarks that it is desirable for an efficient smoothing formula, one that achieves adequate smoothness without producing unnecessarily large departures from the observed values, to have its characteristic function satisfy

$$0 \leqslant \phi(t) \leqslant 1 \tag{6.2}$$

for all t, a stronger condition than (4.3) or (4.4). This remark seems to have been little noted in the years since its publication. We shall call an MWA a *smoothing formula in the strict sense* if its characteristic function satisfies (6.2).

LEMMA 6.1. *Under the natural extension of a given symmetrical* MWA, $D^{-1} - KK^T$ *is nonsingular if and only if* G *is nonsingular, and* H *defined by* (6.1) *is positive definite if and only if* G *is positive definite.*

Proof. If

$$G = I - K^T DK \tag{6.3}$$

as in (1.3), then by Noble's theorem

$$G^{-1} = I + K^T(D^{-1} - KK^T)^{-1} K, \tag{6.4}$$

provided G and D are nonsingular. Under the natural extension, D is always nonsingular by property (4). In the proof of Noble's theorem, the nonsingularity of $D^{-1} - KK^T$ is shown to follow from that of G and D. On the other hand, if $D^{-1} - KK^T$ is nonsingular, multiplication of the right members of (6.3) and (6.4) gives the identity. This proves the first statement of the lemma.

Now let H be positive definite. Then, by (6.4) and (6.1),

$$G^{-1} = I + K^T HK.$$

If v is an arbitrary nonzero real vector,

$$v^T G^{-1} v = v^T v + (Kv)^T HKv. \tag{6.5}$$

The second term of the right member of (6.5) is nonnegative, since H is positive definite, and the first term is positive. It follows that G^{-1}, and therefore G, is positive definite.

Conversely, let G be positive definite. Applying Noble's theorem to (6.3) gives

$$H = D + DK(I - K^T DK)^{-1} K^T D = D + DKG^{-1}K^T D.$$

Now, we note that under the natural extension D is positive definite [8, Theorem 1], since all the zeros of $A(x)$ are outside the unit circle. Thus, the same argument used previously shows that $v^T Hv > 0$ for every nonzero real vector v, and so H is positive definite. This completes the proof.

THEOREM 6.2. *Under the natural extension of a given symmetrical* MWA, *H is positive definite for all N if and only if $\phi(t)$ satisfies* (6.2).

Proof. By Lemma 6.1, H is positive definite if and only if G is positive definite; therefore we need only consider the positive definiteness of G. We recall that $G = I - F$, where F is a singular, symmetric Trench matrix charcterized by two identical polynomials equal to $\hat{A}(x)$. Since all the zeros of $\hat{A}(x)$, with the exception of $+1$, are outside the unit circle, F is positive semidefinite [8, Theorem 1], and if

$$f(x) = \hat{A}(x)\hat{A}(1/x),$$

then

$$\psi(t) = f(e^{it}) = |\hat{A}(e^{it})|^2$$

is nonnegative for all t. Let M denote the maximum of $\psi(t)$.

Now, let $\phi(t)$ satisfy (6.2). Since $\phi(t) = 1 - \psi(t)$, it follows that

$$0 \leqslant \psi(t) \leqslant 1 \tag{6.6}$$

for all t. Therefore $M \leqslant 1$, and it follows [8, Theorem 2] that for all N all eigenvalues of F are nonnegative and less than unity. Since the eigenvalues of G are 1 minus those of F, all of the former are positive for all N, and therefore G is positive definite for all N.

Conversely, let G be positive definite for all N. Then all its eigenvalues are positive for all N, and consequently those of F are less than unity (but not less than zero, since F is positive semidefinite). Since M is the limit of the largest eigenvalue as N becomes infinite [8, Theorem 2], $M \leqslant 1$. Therefore (6.6) holds, and it is equivalent to (6.2). This completes the proof.

It is easy to construct an MWA that is a smoothing formula in the strict sense. However, none of the weighted averages in common use fall in this class. As a practical matter, the smoothing effected by such formulas is likely to be too "gentle." In particular, using properties of Jacobi polynomials, I have shown in [6] that the characteristic functions of all the minimum-R_s averages commonly used by actuaries and economic statisticians (see [10]) assume negative values in $(0, 2\pi)$. Thus no such formula is a smoothing formula in the strict sense.

There is, however, one family of moving averages, mentioned in the literature but not in general use, that are smoothing formulas in the strict sense. This is the limiting case of the minimum-R_s formulas as s approaches infinity [6]. In finite-difference form, the minimum-R_s formula of $2m+1$ terms, exact for the degree $2s-1$, is

$$u_x = \mu^{2(m-s+1)} \sum_{j=0}^{s-1} (-4)^j \binom{m-s+j}{j} \delta^{2j} y_x,$$

where the operator μ is defined by

$$\mu f(x) = \tfrac{1}{2}[f(x+\tfrac{1}{2}) + f(x-\tfrac{1}{2})],$$

so that $\mu^2 = 1 + \frac{1}{4}\delta^2$. The characteristic function is

$$\phi(t) = (\cos \tfrac{1}{2}t)^{2(m-s+1)} \sum_{j=0}^{s-1} \binom{m-s+j}{j} \sin^{2j} \tfrac{1}{2}t,$$

which is nonnegative in $0 < t < 2\pi$, with a single zero of multiplicity $2(m-s+1)$ at $t=\pi$.

It may be mentioned that, in the case where $\phi(t)$ assumes some negative values (and G and H are nonsingular), though the loss function does not have an extremum, the natural extension does correspond to a saddle point of that function.

ACKNOWLEDGMENTS

It will be obvious to the reader that the results obtained in this paper depend heavily on the earlier work of I. J. Schoenberg. I also wish to thank W. F. Trench for his substantial contribution. Indeed, the main theorem (Theorem 5.1) was originally his conjecture.

REFERENCES

1. Ø. BORGAN, On the theory of moving average graduation, *Scand. Actuar. J.* (1979), 83–105.

2. A. Ben-Israel and T. N. E. Greville, "Generalized Inverses: Theory and Applications," Wiley, New York, 1974.
3. A. Dresden, On the iteration of linear homogeneous transformations, *Bull. Amer. Math. Soc.* **48** (1942), 577–579.
4. M. D. W. Elphinstone, Summation and some other methods of graduation—the foundations of theory, *Trans. Fac. Actuar.* **20** (1951), 15–77.
5. T. N. E. Greville, On smoothing a finite table: a matrix approach, *J. Soc. Indust. Appl. Math.* **5** (1957), 137–154.
6. T. N. E. Greville, On stability of linear smoothing formulas, *SIAM J. Numer. Anal.* **3** (1966), 157–170.
7. T. N. E. Greville, On a problem concerning band matrices with Toeplitz inverses, *in* "Proceedings 8th Manitoba Conference Numer. Math. Comput.," pp. 275–283, Utilitas Mathematica, Winnipeg, 1979.
8. T. N. E. Greville, Bounds for eigenvalues of Hermitian Trench matrices, *in* "Proceedings 9th Manitoba Conference Numer. Math. Comput.," pp. 241–256, Utilitas Mathematica, Winnipeg, 1980.
9. T. N. E. Greville, Moving-weighted-average smoothing extended to the extremities of the data. I. Theory, *Scand. Actuar. J.* (1981), 38–55.
10. T. N. E. Greville, Moving-weighted-average smoothing extended to the extremities of the data. II. Methods, *Scand. Actuar. J.* (1981), 65–81.
11. T. N. E. Greville and W. F. Trench, Band matrices with Toeplitz inverses, *Linear Algebra Appl.* **27** (1979), 199–209.
12. E. J. Hannan, "Multiple Time Series," Wiley, New York, 1970.
13. B. Noble, "Applied Linear Algebra," Prentice-Hall, Englewood Cliffs, N. J., 1969.
14. R. Oldenburger, Infinite powers of matrices and characteristic roots, *Duke Math. J.* **6** (1940), 357–361.
15. I. J. Schoenberg, Contributions to the problem of approximation of equidistant data by analytic functions, *Quart. Appl. Math.* **4** (1946), 45–99; 112–141.
16. I. J. Schoenberg, Some analytical aspects of the problem of smoothing, *in* "Studies and Essays Presented to R. Courant on His 60th Birthday," pp. 351–370, Interscience, New York, 1948.
17. I. J. Schoenberg, On smoothing operations and their generating functions, *Bull. Amer. Math. Soc.* **59** (1953), 199–230.
18. V. Thomée, Stability of difference schemes in the maximum-norm, *J. Differential Equations* **1** (1965), 273–292.
19. N. Wiener, "Extrapolation, Interpolation, and Smoothing of Stationary Time Series with Engineering Applications," Wiley, New York, 1949.