SIAM J. APPL. MATH.
Vol. 19, No. 2, September 1970

CONTINUOUS SMOOTHING WITH POLYNOMIAL WEIGHTING*

WILLIAM F. TRENCH†

1. Introduction. We consider finite memory continuous filters with polynomial weighting designed to smooth or differentiate a signal consisting of an unknown polynomial of given degree plus stationary noise. The derivatives $W^{(\nu)}, W^{(\nu+1)}, \cdots, W^{(\nu+m-1)}$ ($\nu \geqq 0, m \geqq 1$) of the weighting functions are required to vanish at the endpoints of the smoothing interval; if $\nu = 0$, this gives an advantage over least squares smoothing in that, when defined to be zero outside the smoothing interval, the weighting functions have $m - 1$ continuous derivatives in $(-\infty, \infty)$. If $\nu = 0$ or $\nu = m$, the method discussed here is analogous to the minimum R_m method of discrete smoothing.

It is shown that these weighting functions can be expressed as convex linear combinations of their least squares counterparts, so that previously established estimates for the generating functions of least squares polynomial smoothing [1], [2] also hold for the new methods.

2. Formulation of the problem. Let H_n be the space of polynomials of degree not greater than n and let $f = p + \varepsilon$, where p is in H_n and ε is a sample function from a zero mean stationary random process with spectral density ψ. We consider smoothing operations of the form

$$g(x) = \int_{-1}^{1} W(y)f(x + y)\,dy, \tag{1}$$

where W is chosen so that

$$Eg(x) = \int_{-1}^{1} W(y)p(x + y)\,dy = p^{(r)}(x); \tag{2}$$

here r is a fixed integer ($0 \leqq r \leqq n$) and the equality is to hold for every p in H_n. The output g from (1) consists of (2) plus a zero mean random component

$$\mathrm{E}(x) = \int_{-1}^{1} W(y)\varepsilon(x + y)\,dy$$

with spectral density

$$\Phi(\theta) = |C(\theta)|^2\psi(\theta),$$

where

$$C(\theta) = \int_{-1}^{1} e^{-ix\theta}W(x)\,dx \tag{3}$$

is the generating function of the smoothing operation. The variance of the smoothed

* Received by the editors December 10, 1968.
† Department of Mathematics, Drexel University, Philadelphia, Pennsylvania 19104.

estimate of $p^{(r)}$ is

$$\sigma^2(E) = \int_{-\infty}^{\infty} |C(\theta)|^2 \psi(\theta)\, d\theta. \tag{4}$$

If $\{\varepsilon\}$ is r-times differentiable, $\{\varepsilon^{(r)}\}$ has spectral density

$$\phi(\theta) = \theta^{2r}\psi(\theta)$$

and variance

$$\sigma^2(\varepsilon^{(r)}) = \int_{-\infty}^{\infty} \theta^{2r}\psi(\theta)\, d\theta. \tag{5}$$

Condition (2) is satisfied for every p in H_n if and only if

$$\int_{-1}^{1} W(y)y^s\, dy = r!\delta_{rs}, \qquad 0 \leqq s \leqq n; \tag{6}$$

from this it follows, on expanding (3) about $\theta = 0$, that

$$\lim_{\theta \to 0} C(\theta)/\theta^r = (-i)^r. \tag{7}$$

Schoenberg [3] has given reasons for calling a smoothing method stable (when $r = 0$) if $|C(\theta)| < 1$ for $\theta \neq 0$. We will say that a smoothing method is r-stable if its generating function satisfies (7) and

$$|C(\theta)/\theta^r| < 1, \qquad 0 \neq 0;$$

the desirability of this property is obvious from (4) and (5).

If F is square integrable in $[-1, 1]$, then the polynomial Q in H_n that minimizes

$$I = \int_{-1}^{1} (F(x) - Q(x))^2\, dx$$

is

$$Q(x) = \int_{-1}^{1} K(x, y)F(y)\, dy,$$

where

$$K(x, y) = \sum_{j=0}^{n} \left(j + \frac{1}{2}\right) P_j(x)P_j(y)$$

and P_j is the Legendre polynomial. Taking $F(y) = y^r$ we can easily see that

$$w_{rn}(x) = \sum_{j=0}^{n} \left(j + \frac{1}{2}\right) P_j^{(r)}(0)P_j(x) \tag{8}$$

satisfies (6); smoothing with this weighting function is called *least squares smoothing*. Among all continuous functions that satisfy (6), w_{rn} is the unique one that minimizes

$$J(W) = \int_{-1}^{1} W^2(x)\, dx = \frac{1}{2\pi}\int_{-\infty}^{\infty} |C(\theta)|^2\, d\theta,$$

which yields the known result that smoothing with w_{rn} provides the minimum variance estimate of $p^{(r)}$ (of the form (1)) if $\{\varepsilon\}$ is white noise ($\psi = 1$).

The question of stability of smoothing with w_{0n} was introduced by Wilf [4]. L. Lorch and P. Szego [5], following Wilf's approach, subsequently showed that continuous least squares smoothing is stable for $r = 0$ and n sufficiently large. The author [1], [2] showed that least squares smoothing is r-stable for all r and n. Lorch and Szego [6] gave an independent proof, using a method different from that of [1], [2], of stability for all n when $r = 0$.

In this paper we consider smoothing with the polynomial weighting function of lowest degree that satisfies (6) and the boundary conditions

$$W^{(j)}(1) = W^{(j)}(-1), \qquad \nu \leqq j \leqq \nu + m - 1. \tag{9}$$

In § 4 we show that this smoothing method is r-stable and give a recursive method for obtaining W; in § 5 we show that if $\nu = 0$ or $\nu = m$ this method is an analogue of minimum R_m smoothing of discrete data, which was recently shown to be stable (for $r = 0$) by Greville [7].

If $\nu = 0$ in (9) the continuity of W (defined to be zero for $|x| > 1$) and its first $m - 1$ derivatives in $(-\infty, \infty)$ is a desirable feature if the filter (1) is to be approximated by a physical device, because it avoids the sharp cutoff that would be required to approximate the discontinuities in a similarly extended w_{rn}.

3. Properties of least squares weighting functions. To aid our investigation of the new smoothing methods we develop some useful properties of least squares weighting functions.

Since $P_j(-x) = (-1)^j P_j(x)$, (8) implies that

$$w_{rn}(-x) = (-1)^r w_{rn}(x) \tag{10}$$

and, depending upon the parity of $n - r$, either $w_{r,n-1} = w_{rn}$ or $w_{rn} = w_{r,n+1}$; we will assume without loss of generality that $n - r$ is even, so that $w_{rn} = w_{r,n+1}$ and

$$\int_{-1}^{1} w_{rn}(x)x^s \, dx = r!\delta_{rs}, \qquad 0 \leqq s \leqq n + 1. \tag{11}$$

LEMMA 1. *Any polynomial of the form*

$$G = \sum_{s=0}^{R} K_s w_{r,n+2s}, \tag{12}$$

where $K_0, \cdots, K_R$ are constants and $K_0 \neq 0$, has at least n sign changes in $(-1, 1)$.

Proof. If $r = n$, G is orthogonal to H_{n-1}, which gives the result; if $r = 0$, the same argument applies to $x^2G(x)$. Now suppose $0 < r < n = 2k$. Then

$$\int_{-1}^{1} G(x)\, dx = 0;$$

hence G changes sign at least once in $(-1, 1)$ and, since G is even, it must change sign at points $\pm x_1, \cdots, \pm x_j$, where $0 < x_1 < \cdots < x_j < 1$. Let

$$\begin{aligned} q(x) &= (x^2 - x_1^2) \cdots (x^2 - x_j^2) \\ &= b_0 + b_1x^2 + \cdots + b_jx^{2j}. \end{aligned}$$

If $j < k$ then for $i \geqq k$,

$$\begin{aligned}\int_{-1}^{1} & w_{r,2i}(x)q(x)(b_{r-1} - b_r x^2)\,dx \\ &= b_{r-1}\int_{-1}^{1} w_{r,2i}(x)q(x)\,dx - b_r\int_{-1}^{1} w_{r,2i}(x)x^2 q(x)\,dx \\ &= r!(b_{r-1}b_r - b_r b_{r-1}) = 0.\end{aligned}$$

Thus from (12),

$$\int_{-1}^{1} G(x)q(x)(b_{r-1} - b_r x^2)\,dx = 0,$$

which is a contradiction since Gq does not change sign in $(-1, 1)$ and $b_r b_{r-1} < 0$; thus $j \geqq k$. This establishes the theorem if n is even; if $0 < r < n = 2k + 1$ a similar argument shows that $xG(x)$ has at least $2k$ sign changes in $(-1, 1)$ (not at the origin), so that G has at least $2k + 1$.

LEMMA 2.

(13) $$i^{r-s}w_{rn}^{(s)}(0) > 0, \qquad s \equiv r (\mathrm{mod}\ 2), \quad s \leqq n,$$

and

(14) $$w_{rn}^{(s)}(0) = 0, \qquad s \not\equiv r (\mathrm{mod}\ 2), \quad s \leqq n - 1.$$

Proof. Equation (10) implies (14); since Lemma 1 and (10) imply that w_{rn} has n distinct roots symmetrically placed about the origin, the derivatives (13) whose orders have the same parity as r are nonzero and alternate in sign; furthermore,

$$w_{rn}^{(r)}(0) = \int_{-1}^{1} w_{rn}^2(x)\,dx > 0,$$

which completes the proof of (13).

LEMMA 3.

(15) $$i^{n-r}w_{rn}^{(s)}(1) > 0, \qquad 0 \leqq s \leqq n.$$

Proof. (Recall that $n - r$ is even.) The signs of the derivatives near the origin are given by Lemma 2; by Rolle's theorem, $w_{rn}^{(s)}$ has $n - s$ sign changes placed symmetrically about the origin, which gives (15).

4. The new polynomial weighting functions. Let W_{mn} be the polynomial of lowest degree that satisfies

(16) $$\int_{-1}^{1} W(x)x^s\,dx = r!\delta_{rs}, \qquad 0 \leqq s \leqq n,$$

and

(17) $$W^{(j)}(1) = W^{(j)}(-1) = 0, \qquad v \leqq j \leqq v + m - 1.$$

Clearly W_{mn} also depends upon r and v, but we will not indicate this by additional subscripts; r and v are fixed in the following discussion. We assume that $0 \leqq r \leqq n$

(if $r > n$, $W_{mn} \equiv 0$) and that $v \leqq n$ (if $v > n$, $W_{mn} = w_{rn}$). As before, we assume that $n - r$ is even.

LEMMA 4. *There is at most one polynomial of degree less than* $n + 2m + 1$ *that satisfies* (16) *and* (17).

Proof. If W_{mn} and $\hat{W}_{mn}$ both satisfy (16) and (17), then $h = W_{mn} - \hat{W}_{mn}$ is orthogonal to H_n and therefore has at least $n + 1$ distinct zeros in $(-1, 1)$, so that $h^{(v)}$ has at least $n - v + 2m + 1$ zeros (counting multiplicities at $x = \pm 1$) in $[-1, 1]$; thus either $\deg h > n + 2m$ or $h \equiv 0$.

THEOREM 1. *If* $0 \leqq r, v \leqq n$, $m \geqq 1$, *and* $n - r$ *is even, then* W_{mn} *is unique and of exact degree* $n + 2m$. *Furthermore*,

$$W_{mn} = \sum_{j=0}^{m} A_{jmn} w_{r,n+2j}, \tag{18}$$

where $A_{0mn}, \cdots, A_{mmn}$ *are constants and*

$$A_{jmn} > 0, \tag{19a}$$

$$\sum_{j=0}^{m} A_{jmn} = 1. \tag{19b}$$

Proof. We use induction on m. For $m = 1$,

$$W_{1n} = \frac{w_{r,n+2}^{(v)}(1) w_{rn} - w_{rn}^{(v)}(1) w_{r,n+2}}{w_{r,n+2}^{(v)}(1) - w_{rn}^{(v)}(1)}; \tag{20}$$

from (15) the denominator is nonzero and the coefficients of w_{rn} and $w_{r,n+2}$ satisfy (19). W_{1n} satisfies (16) because of (11) and $W_{1n}^{(v)}(1) = W_{1n}^{(v)}(-1) = 0$; the uniqueness follows from Lemma 4.

Now suppose the theorem is true for some $m \geqq 1$. Then any weighted average of W_{mn} and $W_{m,n+2}$ satisfies (16) and (17); we have only to choose coefficients so that

$$W_{m+1,n}^{(v+m)}(1) = W_{m+1,n}^{(v+m)}(1) = 0.$$

Hence

$$W_{m+1,n} = \frac{W_{mn}^{(v+m)}(1) W_{m,n+2} - W_{m,n+2}^{(v+m)}(1) W_{mn}}{W_{mn}^{(v+m)}(1) - W_{m,n+2}^{(v+m)}(1)}. \tag{21}$$

To complete the induction we need only show that

$$W_{mn}^{(v+m)}(1) W_{m,n+2}^{(v+m)}(1) < 0. \tag{22}$$

From (17),

$$W_{mn}^{(v)}(x) = (x^2 - 1)^m Q_{mn}(x),$$

where Q_{mn} is of degree $n - v$ and, by Lemma 1 and Rolle's theorem, has exactly $n - v$ sign changes in $(-1, 1)$. Hence $W_{m,n+2}^{(v)}$ has exactly two more sign changes in $(-1, 1)$ than $W_{mn}^{(v)}$; since both are either even or odd we conclude that the former has exactly one more sign change in $(0, 1)$ than the latter. Both are convex combinations of $w_{rn}^{(v)}, w_{r,n+2}^{(v)}, \cdots, w_{r,n+2m+2}^{(v)}$ (by the induction assumption) and, by

Lemma 2, have the same sign near the origin. Thus $Q_{mn}(1)Q_{m,n+2}(1) < 0$ and since

$$W_{mn}^{(v+m)}(1) = 2^m m! Q_{mn}(1),$$

(22) is established and the proof of Theorem 1 is complete.

Equations (8), (20) and (21) provide a recursive method for obtaining W_{mn}; it can also be written explicitly as

$$W_{mn} = \frac{1}{D} \begin{vmatrix} w_{rn} & w_{r,n+2} & \cdots & w_{r,n+2m} \\ a_{00} & a_{01} & \cdots & a_{0m} \\ a_{10} & a_{11} & \cdots & a_{1m} \\ \vdots & \vdots & & \vdots \\ a_{m-1,0} & a_{m-1,1} & & a_{m-1,m} \end{vmatrix}, \tag{23}$$

where

$$a_{ij} = w_{r,n+2j}^{(v+i)}(1), \qquad 0 \leqq j \leqq m, \quad 0 \leqq i \leqq m-1,$$

and

$$D = \begin{vmatrix} 1 & 1 & \cdots & 1 \\ a_{00} & a_{01} & \cdots & a_{0m} \\ a_{10} & a_{11} & \cdots & a_{1m} \\ \vdots & \vdots & & \vdots \\ a_{m-1,0} & a_{m-1,1} & \cdots & a_{m-1,m} \end{vmatrix}.$$

Expansion of the determinant in (23) by elements of the first row yields (18) and (19b); however, the proof of (19a) is more straightforward by the induction argument. Theorem 1 ensures that $D \neq 0$ if $v \leqq n$.

Our main result now follows easily.

THEOREM 2. *The smoothing method* (1) *with* $W = W_{mn}$ *is r-stable.*

Proof. The generating function of W_{mn} is a convex combination of the generating functions

$$C_{rk}(\theta) = \int_{-1}^{1} w_{rn}(x) e^{-ix\theta}\, dx, \qquad k = n, n+2, \cdots, n+2m,$$

which are known to be r-stable [1], [2].

5. Connection with minimum R_m smoothing. If $v = 0$ or $v = m$, W_{mn} is the solution of an extremal problem similar to that of minimum R_m smoothing of discrete data.

THEOREM 3. W_{mn} *(with* $v = 0$ *in* (17)*) is the unique function in* $C^{(2m)}[-1, 1]$ *that minimizes*

$$Q = \int_{-1}^{1} (W^{(m)}(x))^2\, dx$$

subject to (6) *and the endpoint conditions*

$$W^{(j)}(1) = W^{(j)}(-1) = 0, \qquad 0 \leqq j \leqq m-1.$$

If $m \leqq n$, $W = W_{mn}$ (*with* $v = m$ *in* (17)) *is the unique function in* $C^{(2m)}[-1, 1]$ *that minimizes* Q *subject to* (6) *only* (*no endpoint conditions*); *if* $m > n$, *any linear combination of* $w_{rn}, \cdots, w_{r,m-1}$ *is a solution of this problem with* $Q = 0$.

Proof. Both extremal problems are of the isoperimetric type; hence there exist unique constants $\lambda_0, \lambda_1, \cdots, \lambda_n$ such that the solutions of the given problems are extremals of the functional

$$J(W) = \int_{-1}^{1} \left[(W^{(m)}(x))^2 - 2W(x) \sum_{j=0}^{n} \lambda_j x^j \right] dx$$

whose first variation is given by

$$\begin{aligned} \delta J[W; h] &= 2 \int_{-1}^{1} \left(W^{(m)}(x) h^{(m)}(x) - h(x) \sum_{j=0}^{n} \lambda_j x^j \right) dx \\ &= 2 \sum_{j=0}^{m-1} (-1)^j W^{(m+j)}(x) h^{(m-j-1)}(x) \Big|_{-1}^{1} \\ &\quad + 2 \int_{-1}^{1} \left[(-1)^m W^{(2m)}(x) - \sum_{j=0}^{n} \lambda_j x^j \right] h(x)\, dx. \end{aligned} \tag{24}$$

If W is an extremal for the first problem of the theorem, then $\delta J[W; h] = 0$ for every m-times differentiable h such that $h^{(j)}(1) = h^{(j)}(-1) = 0$, $0 \leqq j \leqq m-1$; hence,

$$W^{(2m)}(x) = (-1)^m \sum_{j=0}^{n} \lambda_j x^j \tag{25}$$

and W is a polynomial of degree not exceeding $n + 2m$; by the uniqueness proved in Theorem 1, $W = W_{mn}$ with $v = 0$.

For the second problem, $\delta J[W; h] = 0$ for every $2m$-times differentiable h; hence (25) again applies. The sum on the right side of (24) must vanish for every h in $C^{(2m)}[-1, 1]$, which implies that W satisfies the natural endpoint conditions

$$W^{(j)}(1) = W^{(j)}(-1) = 0, \qquad m \leqq j \leqq 2m-1.$$

If $m \leqq n$, our conclusion follows from Theorem 1, while if $m > n$, the conclusion is obvious.

REFERENCES

[1] W. F. TRENCH, *On the stability of midpoint smoothing with Legendre polynomials*, Proc. Amer. Math. Soc., 18 (1966), pp. 191–199.

[2] ———, *Bounds on the generating functions of certain smoothing operations*, Ibid., 18 (1966), pp. 200–206.

[3] I. J. SCHOENBERG, *On smoothing operations and their generating functions*, Bull. Amer. Math. Soc., 59 (1953), pp. 199–230.

[4] H. S. WILF, *The stability of smoothing by least squares*, Proc. Amer. Math. Soc., 15 (1964), pp. 933–937; Errata, Ibid., 17 (1966), p. 542.

[5] L. LORCH AND P. SZEGO, *A Bessel function inequality connected with the stability of least square smoothing*, Ibid., 17 (1966), pp. 330–332.

[6] ———, *A Bessel function inequality connected with stability of least square smoothing. II*, Glasgow Math. J., 9 (1968), pp. 119–122.

[7] T. N. E. GREVILLE, *On the stability of linear smoothing formulas*, SIAM J. Numer. Anal., 3 (1966), pp. 157–170.