The degree of copositive approximation and a computer algorithm

Yingkang Hu
Xiang Ming Yu
Missouri State University

Follow this and additional works at: https://bearworks.missouristate.edu/articles-cnas

Recommended Citation

This article or document was made available through BearWorks, the institutional repository of Missouri State University. The work contained in it may be protected by copyright and require permission of the copyright holder for reuse or redistribution. For more information, please contact bearworks@missouristate.edu.
The Degree of Copositive Approximation and a Computer Algorithm

YINGKANG HU$^1$ and XIANG MENG YU$^1$

Abstract. Our main results are: (1) Let $f \in C[0, 1]$ change its sign a finite number of times, then the degree of copositive approximation of $f$ by splines with $n$ equally spaced knots is bounded by $C \omega(f, 1/n)$ for an $n$ large enough. This rate is the best in the sense that $C \omega(f, 1/n)$ cannot be replaced by $\omega(f, 1/n)$; (2) an algorithm is developed based on the proof; (3) the first result above holds for copositive polynomial approximation of $f$; (4) if $f \in C^2[0, 1]$, then the degree of approximation by copositive splines of order $r$ is bounded by $C \omega^{-1}(f, 1/n)$.

The results on $f \in C[0, 1]$ fill a gap left by S. P. Zhou [10], and Y. K. Hu, D. Leviatan and X. M. Yu [4,5].

Key words. Computer algorithm, Degree of copositive approximation, Spline approximation, Polynomial approximation, Splines with equally spaced knots

AMS subject classifications. 41-04, 41A10, 41A15, 41A17, 41A25, 41A29

1. Introduction and Main Results

Let $C^r[0, 1]$ be the space of $r$ times continuously differentiable functions on $[0, 1]$, and let $C[0, 1]$ be the space of continuous functions. Let $f \in C[0, 1]$, then we denote by $\| \cdot \|$ the uniform norm taken over $[0, 1]$ and by $\omega_r(f, t)$ the usual $r$th modulus of smoothness of $f$, with $\omega_0(f, t)$ understood as $\|f\|$. We say that $f(x)$ changes sign at $y \in (0, 1)$ if: (1) $f(y) = 0$; (2) there exists an $\varepsilon_1 > 0$ such that $\eta f(x) > 0$ for any $x \in (y - \varepsilon_1, y)$, where $\eta = \pm 1$; (3) there exists an $\varepsilon_2 > 0$ such that $\eta f(x) < 0$, for any $x \in (y, y + \varepsilon_2)$; (4) the inequalities hold strictly at least at one $x$ in (therefore on a portion of) each of the neighborhoods above, that is, $f$ has a true sign change at $y$. Such a $y$ is called a point of sign change of $f$. We assume $f$ only has $k < \infty$ sign changes at $0 < y_1 < y_2 < \cdots < y_k < 1$, and denote $y_0 := 0$ and $y_{k+1} := 1$. A function $g$ is said to be copositive with $f$ if it has the same sign with $f$ on each interval $[y_i, y_{i+1})$ and changes its sign exactly at each $y_i$.

For any $n > 0$, let $h := 1/n$, $x_i := ih$ and $T := \{x_i\}$, then $\{N_{i,r}\}_{i=1,r}$ forms a basis for all splines of order $r$ on $[0, 1]$ on knot sequence $T$, where

\[ N_{i,r}(x) := (x_{i+r} - x_i)[x_i, \ldots, x_i] (x - x_i)^{(r-1)}_+ \]

are the B-splines of order $r$ on $T$, and every such a spline can be written as

\[ s(x) = \sum_{i=1}^{n-1} c_i N_{i,r}(x), \]

Since all the knots are simple, $s \in W^{r-1}_\infty$, here and throughout $W^k_p$ denotes the Sobolev space, the space of functions whose $(k - 1)$st derivative is absolutely continuous and whose $k$th derivative lies in $L_p$, $1 \leq p \leq \infty$. 
The purpose of this paper is to investigate the degree of approximation of functions $f \in C[0, 1]$ by splines and polynomials that are copositive with $f$. Although the degree of nonnegative approximation to a nonnegative function is of the same order as the best unconstrained approximation, it is not so for copositive approximation. For example, S. P. Zhou [10] proves that it is impossible to approximate a general function $f \in C[0, 1]$ by copositive polynomials at a rate of $\omega_4$ even if this $f$ has a continuous derivative and changes its sign only once in $[0, 1]$. 

**Theorem A.** There is a function $f \in C^1[0, 1]$ which changes its sign once in $[0, 1]$, such that

$$
\limsup_{n \to \infty} \frac{E_n^{(1)}(f)}{\omega_4(f, 1/n)} = +\infty,
$$

where $E_n^{(1)}(f)$ is the error of the best copositive approximation to $f$ by polynomials of degree $n$. 

Y. K. Hu, D. Leviatan and X. M. Yu [5] have recently proved that this also holds for copositive approximation by splines with equally spaced knots, as shown in the following theorem. 

**Theorem B.** There is a function $f \in C^1[0, 1]$ which changes its sign once in $[0, 1]$, such that for any sequence of splines $s_n$ of order $r$ with $n-1$ equally spaced interior knots, which are copositive with $f$, we have

$$
\limsup_{n \to \infty} \frac{\|f - s_n\|}{\omega_4(f, 1/n)} = \infty.
$$

In fact, this type of constrained approximation is impossible even for a much larger class of approximating functions [5, Theorem 4]. On the positive side, they prove in the same paper that, under conditions very similar to those in our main results, the degree of copositive approximation by either polynomials or splines with equally spaced knots is no worse than $\omega_2$. Their approach is to construct the desired spline first, and then apply the theorem below to the spline to get the polynomial approximant. We will use the same approach to prove one of our theorems (Theorem 2).

**Theorem C.** Let $f \in C^1[0, 1]$ change its sign $k \geq 1$ times at $0 < y_1 < y_2 < \cdots < y_k < 1$ and let $\delta := \min_{0 < j < k} (y_{j+1} - y_j)$, where $y_0 := 0$ and $y_{k+1} := 1$. Then for any $m \geq 1$, there are positive constants $C_1 = C_1(k, m)$ and $C_2 = C_2(k, m, \delta)$, such that for each $n > C_1 \delta^{-1}$, there exists a polynomial $p_n$ of degree $n$ which is copositive with $f$ and satisfies

$$
\|f - p_n\| \leq C_2 n^{-1} \omega_m(f, 1/n).
$$

The results in [5] and [10] leave an obvious gap of $\omega_3$. The reason Hu, Leviatan and Yu can not get order $\omega_3$ is that the construction of the spline in [5] depends on the Schoenberg-Bernstein variation diminishing operator, which preserves many geometric properties of $f$ but only has a degree of approximation $\omega_2$. Theorem D and Corollary E below, which are very recent results of ours [6], enable us to start the construction with a best quadratic spline approximation of $f$, which yields degree $\omega_3$.

**Theorem D.** Let $1 \leq r < n$ and $0 \leq m \leq r$, and let $s$ be as in (1.1). Then

$$
h^j \omega_{m-j}(s^{(j)}, h) \sim h^{1/p} \|\Delta^m c\|_p, \quad 0 \leq j \leq \min(m, r-1),
$$

therefore they are all equivalent to each other with the equivalence constants depending
only on r, where e is the B-spline coefficient sequence of s and $\Delta^n$ is the mth difference operator on sequences.

Corollary E. Let $f \in W^2_{\varepsilon}[0, 1]$, $j \geq 0$, have a spline approximant $s_n$, $n > r$, of the form (1.1) such that

$$\|f - s_n\|_p \leq C n^{-j} \omega_m(f^{(j)}, 1/n)_p \quad \text{for some } m \leq r - j,$$

then

$$\omega_{m+j}(s^{(k)}_n, 1/n)_p \leq C n^{k-j} \omega_m(f^{(j)}, 1/n)_p, \quad 0 \leq k \leq \min(r - 1, m + j),$$

where C is a constant depending only on r.

Our first result is about copositive approximation to functions in $\mathbf{C}[0, 1]$ by splines with equally spaced knots, which fills the gap mentioned above for splines. We postpone the proof until §2.

Theorem 1. Let $f \in \mathbf{C}[0, 1]$ change its sign $k \geq 1$ times at $0 < y_1 < y_2 < \cdots < y_k < 1$. Let $y_0 := 0$, $y_{k+1} := 1$ and $\delta := \min_{0 \leq j \leq k} (y_{j+1} - y_j)$. Then for each $n \geq 4\delta^{-1}$, there exists a quadratic spline $s_n$ copositive with $f$ with $n-1$ equally spaced interior knots such that

$$\|f - s_n\| \leq C \omega_3(f, 1/n),$$

where $C$ is an absolute constant.

Based on the proof of Theorem 1, we develop an algorithm which actually computes (the coefficients of) the spline in the theorem. It will be described in §3.

With Theorem 1, we can easily prove a similar result on polynomials in the same way as that of [5], and this fills the gap for polynomials.

Theorem 2. Let $f \in \mathbf{C}[0, 1]$ change its sign $k \geq 1$ times at $0 < y_1 < y_2 < \cdots < y_k < 1$. Let $\delta := \min_{0 \leq j \leq k} (y_{j+1} - y_j)$, where $y_0 := 0$ and $y_{k+1} := 1$. Then there are positive constants $C_1 = C_1(k)$ and $C_2 = C_2(k, \delta)$ such that for each $n > C_1\delta^{-1}$, there exists a polynomial $p_n$ of degree $\leq n$ which is copositive with $f$ and satisfies

$$\|f - p_n\| \leq C_2 \omega_3(f, 1/n).$$

Proof. In this proof and throughout this paper, the capital letter C denotes a constant that may change its value from one occurrence to another, even in the same line. We apply Theorem C to the spline $s_n$ in (1.7) and obtain a polynomial $p_n$ of degree $\leq n$ that is copositive with $s_n$, therefore with $f$ as well. One can readily estimate the distance between $p_n$ and $f$:

$$\|f - p_n\| \leq \|f - s_n\| + \|s_n - p_n\| \leq C \omega_3(f, 1/n) + C n^{k-1} \omega_2(s^{(k)}_n, 1/n) \leq C \omega_3(f, 1/n),$$

where (1.4) and (1.6) have been used. $\blacksquare$

Now that $\omega_3$ is the best one can do with general functions in $\mathbf{C}$, a question that arises naturally is whether one can do better with smoother functions. In this aspect, we have the following analogue to Theorem C for splines that says the optimal order of approximation can be achieved in this case. The proof will be given in §4 at the end of this paper.

Theorem 3. Let $f \in C^1[0, 1]$ change its sign $k \geq 1$ times at $0 < y_1 < y_2 < \cdots < y_k < 1$. Let $\delta := \min_{k \leq j \leq k-1} (y_{j+1} - y_j)$ if $k > 1$ and $\delta := 3$ if $k = 1$. Then for any $r \geq 2$ and $n \geq 3\delta^{-1}$, there exists a spline $s_n$ of order $r$ with no more than $n + 4k(r - 1)^2$ interior knots that is copositive with $f$ and satisfies

$$\|f - s_n\| \leq C r n^{((r-1))} \omega_{r-1}(f', 1/n),$$
where $C_r$ is a constant depending only on $r$.

Remark. Although the knots of this $s_n$ are not equally spaced, it can be viewed as a spline on a finer uniform mesh, so Theorem D and Corollary E still apply.

2. Proof of Theorem 1

We define

$$
(2.1) \quad \bar{x}_i := \frac{x_i + x_{i+1}}{2} = x_i + \frac{h}{2}, \quad i = -2, \ldots, n + 1,
$$

to be averages of the knots in $T$ and define $J_i := [\bar{x}_i, x_{i+1}]$. The following simple properties about the B-splines $N_{i,3}$ on $T$ can be easily derived from formulae in almost any book discussing B-splines, for examples, [2] and [8].

(2.2.1) $N_{i,3}$ is symmetric about the line $x = \bar{x}_{i+1}$, increasing on the left and decreasing on the right; it assumes its maximum value $3/4$ at the center $\bar{x}_{i+1}$ of its support $[x_i, x_{i+3}]$; It is convex on $J_i$ and $J_{i+2}$ and concave on $J_{i+1}$.

(2.2.2) Denote $n \sum_{i=-2}^{n-1} c_i N_{i,3}$ by $S$, then $S(x_i) = \frac{c_{i-1} + c_{i+2}}{2}$, $i = 0, \ldots, n$.

(2.2.3) $S' = \sum_{i=-1}^{n-1} \frac{c_i - c_{i-1}}{h} N_{i,2}$ and $S'(x_i) = \frac{c_{i-1} - c_{i+2}}{h}$, $i = 0, \ldots, n$.

The points $P_i(\bar{x}_i, c_{i-1})$, $i = -1, \ldots, n$, totally determine the graph of $S$ and are called the control points of $S$ [2]. In terms of the control points, (2.2.2) and (2.2.3) can be geometrically restated as: the line segment $P_{i-1}P_i$ is tangent to the graph of $S$ at $x = x_i$. We are now ready to prove Theorem 1.

Proof of Theorem 1. We denote $\omega_3(f, 1/n)$ by $\omega_3$ for convenience and fix $n \geq 4/\delta$.

Let

$$
\tilde{s}(x) = \sum_{i=-2}^{n-1} c_i N_{i,3}(x)
$$

be a best quadratic spline approximant of $f$ on $T$, then $\tilde{s}$ satisfies (1.7), namely,

$$
(2.3) \quad E := \| f - \tilde{s} \| \leq C \omega_3.
$$

This $\tilde{s}$ is not copositive with $f$ in general, of course. we have to modify it to get the desired spline $s$.

Before the modification we make two assumptions about $\tilde{s}$. Denote

$$
I_i := [\bar{x}_i, \bar{x}_{i+1}], \quad i = -1, \ldots, n - 1.
$$

We call $I_i$ contaminated if $\bar{x}_i < y_j < \bar{x}_{i+1}$ for some $j$. Since $n \geq 4/\delta$, there is exactly one $y_j$ in each contaminated interval. Let $I_r$ and $I_l$ be any two consecutive contaminated intervals, then

$$
(2.4) \quad 3 \leq r < r + 4 \leq l \leq n - 5,
$$

that is, the distance between them is at least $3h$. Our Assumption 1 is that $\tilde{s}$ has the same sign as $f$ on $[x_{r+2}, x_l]$. Otherwise, we lift or lower the control points $P_i(\bar{x}_i, c_{i-1})$, $i = r + 1, \ldots, l$, by $E$ units. More precisely, if $f$ is nonnegative on $[x_{r+2}, x_l]$, for
example, then for any \( x \) in that interval we have
\[
0 \leq \tilde{s}(x) - f(x) + E \leq \tilde{s}(x) + E
\]
(2.5)
\[
= \sum_{i=r}^{l-1} c_i N_{i,3}(x) + E = \sum_{i=r}^{l-1} (c_i + E) N_{i,3}(x).
\]

If \( \tilde{s} \) is negative in (part of) the interval, we replace \( c_i \) by \( c_i + E, i = r, \ldots, l - 1 \), then \( \tilde{s} \) will satisfy the assumption and, by the definition of \( E \), the inequality in (2.3) will still hold. Similarly, we assume \( \tilde{s} \) has the same sign as \( f \) on \([x_r, x_s]\) if \( l_r \) is the first contaminated interval, and on \([x_{r+2}, x_{n}]\) if it is the last. This means \( \tilde{s} \) may have wrong sign only in the interval \([x_r, x_{r+2}]\) for each contaminated interval \( l_r \).

Our Assumption 2 about \( \tilde{s} \) is that
(2.6)
\[
c_{r-1} \geq c_r \quad \text{or} \quad c_{r-1} \leq c_r
\]
if \( f \) changes sign in \( l_r \), from nonnegative to nonpositive, (or from nonpositive to nonnegative, respectively). Otherwise, we set them both to their average, which would increase (decrease) \( c_{r-1} \), consistent with Assumption 1. This is possible because if \( f \) changes sign from nonnegative to nonpositive and \( c_{r-1} < c_r \), for example, then we will have
\[
0 < -4(c_{r-1} - c_r) \leq 2\tilde{s}(x_r) - 4(c_{r-1} - c_r) = 2\tilde{s}(x_{r+2})
\]
\[
= c_{r-2} + c_{r-1} - 4c_{r-1} + 4c_r - c_r - c_{r+1}
\]
\[
= c_{r-2} - 3c_{r-1} + 3c_r - c_{r+1} \leq C\omega_3(\tilde{s}, 1/n) \leq C\omega_3.
\]

Here Assumption 1 about \( \tilde{s} \) has been used in the second step, (2.2.2) in the third, Theorem D in the fifth and Corollary E in the last one.

We are now ready to modify \( \tilde{s} \). Let \( y_j \) be any of the points of sign changes, and \( l_r \) be the contaminated interval containing \( y_j \). We suppose \( f \) changes sign from nonnegative to nonpositive at \( y_j \), and \( \tilde{s}(y_j) \leq 0 \). The other cases can be proved by replacing \( f(x) \) by \(-f(1-x)\) or \(-f(x)\). To correct the sign of \( \tilde{s} \) near \( y_j \), we add to it a correction function
(2.7)
\[
m_j(x) := -\frac{-\tilde{s}(y_j)}{N_{r-1,3}(y_j)} N_{r-1,3}(x).
\]

Since \( N_{r-1,3}(y_j) \geq 1/8 \) by (2.2.1), and
\[
|\tilde{s}(y_j)| \leq |f(y_j) - \tilde{s}(y_j)| \leq C\omega_3,
\]
we have
(2.8)
\[
\|m_j\| \leq C\omega_3.
\]

As mentioned above, all cases will be changed to this case by rotation or reflection, hence one can see that the support of any correction function contains the corresponding contaminated interval in its interior and extends beyond it by no more than \( 1.5h \) on each side, therefore, by (2.4), the interiors of the supports of the correction functions \( m_j, j = 1, \ldots, k \), will be mutually disjoint, which enables us to define \( s \) by
\[
s(x) = \begin{cases} \tilde{s}(x) + m_j(x), & \text{if } x \text{ is in the support of some } m_j \\ \tilde{s}(x), & \text{otherwise} \end{cases}
\]

Note (1.7) now follows from (2.3) and (2.8).
The only thing remaining to show is that this \( s \) is copositive with \( f \) in \([x_{r-1}, x_{r+2}]\), the support of \( m_j \). This part of the proof will be almost identical to Case 4 in the proof of [5, Theorem 1]. We will take full advantage of the fact that \( s \) is a parabola on each interval \( J_r = [\xi_i, \xi_{i+1}] \). It is obvious that \( s \) is copositive with \( f \) on \( J_{r-1} \) since \( \bar{s}, m_j \) and \( f \) are all nonnegative. On \( J_{r+1} \), we claim that \( s = \bar{s} + m_j \) is either convex or decreasing. We merely need to show this for \( \bar{s} \) instead, because \( m_j \) is both convex and decreasing on \( J_{r+1} \). Indeed, since \( \bar{s}' \) is linear on this interval, \( \bar{s} \) is coconvex with the broken line \( \bar{P}_{r}P_{r+1}P_{r+2} \) by (2.2.3), where \( P_r(\xi_i, c_{i-1}) \), \( i = r, r + 1 \) and \( r + 2 \), are control points defined at the beginning of the section. If \( P_rP_{r+1}P_{r+2} \) is concave, then \( P_{r+1}P_{r+2} \) has smaller slope than that of \( P_rP_{r+1} \), which is already nonpositive by (2.6), and this means \( \bar{s}' \) is nonpositive at both endpoints \( x_{r+1} \) and \( x_{r+2} \) hence on the whole interval, therefore the claim is true.

On \( J_r \) we claim that \( s \) is either concave or decreasing. An argument similar to the above shows this is true on the interval \([\xi_r, x_{r+1}]\). If \( s \) is convex on \( J_r \), then it must be decreasing on \([\xi_r, x_{r+1}] \subset J_r \), therefore it is decreasing on whole \( J_r \).

Now if \( x_{r+1} \leq y \leq x_{r+1} \), then these two claims together with \( s(x_r) \geq 0 \),

\[
s(y) = \bar{s}(y) + m_j(y) = \bar{s}(y) - \frac{\bar{s}(y_j)}{N(y_j)} N(y) = 0
\]

and \( s(x_{r+2}) \leq 0 \) (see Assumption 1) imply that \( s \) is copositive with \( f \) on \( J_{r+1} \) and is non-negative on \( J_r \). If \( \xi_r < y < x_{r+1} \), then the above imply that it is copositive with \( f \) on \( J_r \) and is nonpositive on \( J_{r+1} \).

3. An Algorithm

We have developed an algorithm for copositive approximation based on the proof of Theorem 1, and implemented it on the computer. The code is written in ANSI C. The function \( f \) is supplied in a subprogram as a formula, but only the values at \( x_i \) and \( \xi_i \) are really needed. The input data consists of the end points of an interval \([a, b]\) (instead of \([0, 1]\) ), \( k \) and \( \{y_i\} \). If the given \( k \) is too small, it will be increased to the first integer \( \geq 4(b-a)/\delta \), where \( \delta := \min(1.6(y_n - y_0), y_n - y_0, \ldots, y_k - y_{k-1}, 1.6(y_n - y_{k+1})) \). The factor 1.6 is used here since we only need \( y_n - y_0 \) and \( y_{k+1} - y_k \geq 2.5h \) instead of 4\( h \), see the discussion after (2.8) in §2.

Since it is not easy, and not necessary, to find a best \( L_\infty \) spline approximant for \( \bar{s} \), we use instead the quasi-interpolation scheme described in [8, §6.4], written as a subprogram called \textit{quasi interp}. While the proof uses all simple knots inside and outside the interval \([a, b]\), the algorithm uses multiple knots \( x_{-2} = x_{-1} = x_0 = a \) and \( x_n = x_{n+1} = x_{n+2} = b \) at the ends of the interval, otherwise the quasi-interpolation scheme would require values of \( f \) outside \([a, b]\) where it may not be defined. We point out that any spline approximation scheme, such as linear least squares method, can be used for \( \bar{s} \). In that case, the error of the final spline \( s \) will be bounded by \( C \max(\|f - \bar{s}\|, \omega_3(\bar{s}, (b-a)/n)) \), (see §2).

The modification is done in a subprogram called \textit{modify spline}, which checks the two assumptions about \( \bar{s} \), changes some coefficients if they are not met, then adds the correction functions \( m_j \) to \( \bar{s} \) by changing appropriate coefficients, namely, \( c_r \) or \( c_{r-1} \), depending on the trend of \( f \) near \( y_i \) and the sign of \( \bar{s}(y_i) \), as described in the proof. Since \( 0 \leq N_{i,3} \leq 3/4 \), we merely need to make sure the changes in coefficients are not too large.

While the algorithm follows the proof closely elsewhere, it accomplishes the work involved with Assumption 1 in a totally different way. This is because (2.5) would
make the error unnecessarily large, let alone the cost of estimating the uniform norm of $f - \delta$. ModifySpline checks the sign of the coefficients, and changes them only when it is absolutely necessary. We only describe the case that $I_p$ and $I_q$ are two consecutive contaminated intervals and $f$ is nonnegative between them. The other cases are treated similarly. The program checks the sign of $c_i$, $i = r, r + 1, \ldots, l - 1$, and groups consecutive nonpositive $c_i$'s. If such a group consists of more than one $c_i$, it changes them all to zero. This is justified by the fact that if $c_i, c_{i+1} \leq 0$, then

$$0 \geq \min(c_i, c_{i+1}) \geq c_i + c_{i+1} = 2\delta(x_{i+2}) \geq 2(\delta(x_{i+2}) - f(x_{i+2})) \geq -2 ||f - \delta||.$$  

If a group consists of a single nonpositive $c_i$, modifySpline does one of the following:

1. If $i = r$ and $|c_i| > |c_{i+1}|$, set $c_i$ to $-c_{i+1}$;
2. If $i = l - 1$ and $|c_i| > |c_{i-1}|$, set $c_i$ to $-c_{i-1}$;
3. If $r < i < l - 1$, $c_i < 0$ and $\frac{c_{i-1} + c_i}{2} - \frac{(c_{i-1} - c_i)^2}{2(c_{i+1} - 2c_i + c_{i-1})} < 0$, then set

$$c_i \leftarrow \frac{c_{i-1} + c_{i+1} - \sqrt{c_{i-1}^2 + 14c_{i-1}c_{i+1} + c_{i+1}^2}}{6};$$

4. Otherwise, do nothing.

The change in (1) guarantees that the new $\delta$ satisfies $\delta(x_{i+3}) = 0$, (the value $\delta(x_{i+3}) \geq 0$ is unchanged). By arguments similar to those in §2, the new $\delta$ is nonnegative on $[x_{i+3}, x_{i+4}]$. And inequalities similar to (3.1) show that the change in $c_i$, thus in $\delta$ itself, is less than $C ||f - \delta||$. The change in (2) is the counterpart of this at the other end of the group. The little lemma below, together with (2.2), guarantees that $\delta$ is nonnegative on $[x_{i+1}, x_{i+2}]$ after the change in (3). That it is nonnegative on $[x_i, x_{i+3}]$, the support of $N_{i,3}$, can be shown in a similar way to that of the proof of Theorem 1, provided $\delta(x_i)$ and $\delta(x_{i+3})$ are both nonnegative, which will hold true after modifySpline finishes the processing.

**Lemma 3.1.** Let $P_1(-h/2, c_1)$, $P_2(h/2, c_2)$ and $P_3(3h/2, c_3)$ be three points in the plane with $c_1, c_3 \geq 0$ and $c_2 < 0$ (or $c_1, c_3 < 0$ and $c_2 > 0$), and let $p$ be the quadratic polynomial tangent to the line segments $P_1P_2$ and $P_2P_3$ at their midpoints $(0, (c_1 + c_2)/2)$ and $(h, (c_2 + c_3)/2)$, respectively. Then $p$ assumes its minimum (or maximum) value at $\ddot{x} := \frac{(c_1 - c_2)h}{c_3 - 2c_2 + c_1} \in [0, h]$. This extremum $p(\ddot{x}) = 0$ if

$$c_2 = \frac{c_1 + c_3 - \sqrt{c_1^2 + 14c_1c_3 + c_3^2}}{6} \quad \text{or} \quad \frac{c_1 + c_3 + \sqrt{c_1^2 + 14c_1c_3 + c_3^2}}{6}.$$  

**Proof.** Without loss of generality, we assume $c_1, c_3 \geq 0$ and $c_2 < 0$. One can readily check

$$p(x) = \frac{c_1 + c_2}{2} + \frac{c_2 - c_1}{h}x + \frac{c_3 - 2c_2 + c_1}{2h^2}x^2$$

and $p'' > 0$. Solving

$$p(\ddot{x}) = \frac{c_1 + c_2}{2} - \frac{(c_1 - c_2)^2}{2(c_3 - 2c_2 + c_1)} = 0$$

for $c_2$ gives us one negative value

$$c_2 = \frac{c_1 + c_3 - \sqrt{c_1^2 + 14c_1c_3 + c_3^2}}{6}.$$  

We now show that the change in \((3)\) will not change \(c_i\) too much, (in fact, it is just enough to increase the minimum to zero). Let \(\tilde{x} := \frac{(c_{i-1} - c_i)k}{c_{i+1} - 2c_i + c_{i-1}}\) with \(c_i\) being the new value computed from \((3.2)\). Let \(\tilde{s}_{\text{old}}\) be the value of \(\tilde{s}\) at \(\tilde{x}\) before the change of \(c_i\) and \(\tilde{s}_{\text{new}}\) the value after that. We observe \(c_i\) has been increased, since the old minimum of \(\tilde{s}\) in \([x_{i+1}, x_{i+2}]\) is negative and the new one is \(\tilde{s}_{\text{new}} = 0\) by Lemma 3.1. This means \(\tilde{s}_{\text{old}}\) is negative. Let \(d(c_i)\) be the change in \(c_i\), we have
\[
0 < d(c_i)N_{i,3}(\tilde{x}) = -\tilde{s}_{\text{old}} \leq f(\tilde{x}) - \tilde{s}_{\text{old}}.
\]
Since \(x_{i+1} \leq \tilde{x} \leq x_{i+2}\), we know \(N_{i,3}(\tilde{x}) \geq 1/2\), therefore
\[
0 < d(c_i) \leq 2(f(\tilde{x}) - \tilde{s}_{\text{old}}) \leq C\omega_3.
\]

To evaluate various splines, the program calls the CMLIB (NBS Core Math Library) subprograms DBSPVD and DBVALU, written by C. de Boor and modified by D. E. Amos. The double precision version is used here only because most floating type variables in C are of type \textit{double} by convention. At the end, the program computes values of \(f\), \(\tilde{s}\) and \(s\), and passes them to \textit{Mathematica} (Wolfram Research, Inc.) to produce graphs, including the graphs shown in this paper.

---

**Figure 1.** Approximation of a polynomial with two single zeros and two double ones inside the interval \([-1.2, 1.2]\) with \(n = 12\). Different scales are used for the x- and y-axes.

The algorithm is efficient, because all the computations are local. In fact, it is so fast that we do not feel it is worth CPU timing. On a Sun SPARCstation 2, it took 1–2 seconds elapsed time to approximate a polynomial of degree 8 with \(n = 10\), including evaluation of \(f\), \(\tilde{s}\) and \(s\) at 200 points, the most time consuming part of the program. It doubled with \(n = 190\). As for the order of errors, in all of our tests \(\|f - s\|\) was the same as or slightly larger than \(\|f - \tilde{s}\|\), which is theoretically guaranteed to be of the order \(\omega_3(f, 1/n)\). We give a few figures showing the results. Small values of \(n\) are used so that different curves are distinguishable. In all the figures, \(f\) is represented by solid curve, \(\tilde{s}\) by dotted curve and \(s\) by dashed one.
In preparation of Figure 1, we gave the program a polynomial
\[ f(x) = (x + 1.5)(x + 0.7)(x + 0.55)^2(x - 0.2)^2(x - 0.7)(1.6 - x) \]
on \([-1.2, 1.2]\) with \(n = 10\). This \(n\) turned out to be too small and was changed to 12 by the program. In the figure the final spline \(s\) (dashed) seems different from \(\tilde{s}\) (dotted) only near one of the two single zeros, \(x = -0.7\). Actually, they also differ near \(x = 0.7\), but the difference is too small to be seen, \((0.0026\) at \(x = 0.7\)). The error \(\|f - s\| = 0.03066\) is even smaller than \(\|f - \tilde{s}\| = 0.03039\). When larger values of \(n\) were used in our tests, the two errors were the same for this function.

**Figure 2.** A function that vanishes on the interval \([0.5, 1]\) with a vertical tangent line on each side. The point of sign change is specified as \(y_1 = 0.75\) in the input.

The next figure we want to show is about a function that is identically zero in the middle of the interval \([0, 1.5]\) with two vertical tangent lines:

\[
f(x) = \begin{cases} 
\sqrt{0.5 - x}, & \text{for } 0 \leq x < 0.5 \\
0, & \text{for } 0.5 \leq x \leq 1 \\
-\sqrt{x - 1}, & \text{for } 1 < x \leq 1.5.
\end{cases}
\]

We specified \(y_1 = 0.75\), the midpoint of \([0.5, 1]\) as the point of sign change. Given \(n = 10\), the program produced a final spline \(s\) with \(\|f - s\| = 0.07604\), about 32% larger than \(\|f - \tilde{s}\| = 0.05734\). This \(s\) looks more pleasant than \(\tilde{s}\), preserving much of the shape of the original function \(f\). Note any point in \([0.5, 1]\) can be viewed as the point of sign change of \(f\). We tried \(y_1 = 0.5\) in another test in which \(s\) had an error 0.1061, 85% larger than that of \(\tilde{s}\), and looked worse than \(\tilde{s}\), too.

In Figure 3, the function is basically the same as that in Figure 2, only moved down by 0.01. It then has a unique zero at \(y_1 = 0.4999\). The quasi-interpolant \(\tilde{s}\) is positive near \(x = 0.9\) where \(f\) is negative. Note \(s\) confines itself under the \(x\)-axis. In fact, it is tangent to the \(x\)-axis at \(x = 0.9\). Only part of the graph is shown in the figure so that more details can be seen.
4. Proof of Theorem 3

In this proof, we shall first construct no more than $2k + 1$ splines on overlapping subintervals of $[0, 1]$ with desired properties, then blend them to produce the final spline in the theorem, using the following lemma by R. K. Beatson \cite{1}.

**Lemma 4.1.** Let $r \geq 2$ be an integer and $d = 2(r - 1)^3$. Let $\mathbf{T} = \{x_i\}_{i=-\infty}^{\infty}$ be a strictly increasing knot sequence with $x_0 = a$ and $x_d = b$. Let $g_1, g_2$ be two polynomials of degree $< r$. Then there exists a spline $g$ of order $r$ on $\mathbf{T}$ such that

1. $g(x)$ is a number between $g_1(x)$ and $g_2(x)$ for each $x \in [a, b]$,
2. $g = g_1$ on $(-\infty, a]$ and $g = g_2$ on $[b, \infty)$.

**Proof of Theorem 3.** Let $\mathbf{I}_l := \{x_l, x_{l+1}\}$, where $\{x_l\}$ is the knot sequence $\mathbf{T}$ defined at the beginning of §1. We say $\mathbf{I}_l$ is contaminated if it contains a point $y_j$ of sign change of $f$. Since $n \geq 3/\delta$, there is at least one $y_j$ in each contaminated interval $[\tilde{I}_l, \tilde{I}_{l+1})$, $j = 1, \ldots, k$. Moreover, we have $l_j + 3 \leq l_{j+1}$ for any $j$, that is, the distance between $\tilde{I}_l$ and $\tilde{I}_{l+1}$ is at least $2h$. We approximate $f'$ on $[0, 1]$ by a spline $\hat{s}$ of order $r - 1$ on knot sequence $\{x_l\}$ with

\begin{equation}
E := ||f' - \hat{s}'|| \leq C \omega_{r-1}(f', 1/n).
\end{equation}

Let $\eta_l := 1$ if $f$ changes sign from nonpositive to nonnegative at $y_j$, and $\eta_l := -1$ otherwise. Define on each $\tilde{I}_j := [x_{j-1}, x_{j+2}] \cap [0, 1]$

\begin{equation}
\bar{s}_j(x) := \int_y^x (\hat{s}(t) + \eta_j E) \, dt, \quad j = 1, \ldots, k.
\end{equation}

They are splines of order $r$ on $\{x_l\}$, and are copositive with $f$:

\begin{equation}
\bar{s}_j(y_l) = 0;
\end{equation}
For the degree of error, we have for $x \in J_i$

$$\|f(x) - \tilde{s}_j(x)\| \leq \frac{2}{n}(2\|f' - \tilde{s}'\|) \leq Cn^{-1}\omega_{r-1}(f', 1/n).$$

Denote the closures of the intervals in the complement set $[0, 1] \sim (\cup \tilde{I}_i)$ by $\tilde{I}_i := [x_{i+1}, x_{i+1}], j = 0, 1, \ldots, k$, with $l_0 := -1$ and $l_{k+1} := n$. We observe that on each $\tilde{I}_i$, $f$ does not change its sign, which is given by $\eta_j$. One can readily construct on each $\tilde{I}_i$ a spline $\hat{s}_j$ of order $r$ on $\{x_i\}$ such that

$$\|f - \hat{s}_j\|_{\tilde{I}_i} \leq C\omega_r(f, 1/n) \leq Cn^{-1}\omega_{r-1}(f', 1/n)$$

and

$$\eta_j \hat{s}_j(x) \geq 0, \quad x \in \tilde{I}_i.$$

If $l_1 \leq 1$ (or $l_k \geq n - 2$), then $\hat{s}_0$ (or $\hat{s}_k$) will be unnecessary. In this case we simply skip that value of $j$.

Now that all the overlapping splines have been defined with desired properties, the only thing remaining to do is to blend them to produce the final spline $s \in C^{r-2}[0, 1]$, which is possible by Lemma 4.1. In each admissible pair of $\tilde{s}_{j-1}$ and $\tilde{s}_j$, both of them are polynomials on $[x_{j-1}, x_j]$, the interval where they overlap. We insert a polynomial piece into this interval (by cutting it into $d$ equal pieces), where $d := 2(r - 1)^2$. By Lemma 4.1, there exists a spline $b_j$ consisting of $d$ polynomial pieces whose graph lies between those of $\tilde{s}_{j-1}$ and $\tilde{s}_j$ and connects them in a $C^{r-2}$ manner. We do the same thing to each admissible pair of $\tilde{s}_j$ and $\tilde{s}_k$. Define the final spline to be $b_j$ on these overlapping intervals, and otherwise to be either $\hat{s}_j$ or $\hat{s}_k$, whichever is defined there. It is easy to see from the construction that $s$ is copositive with $f$ and satisfies (1.9), and that $s$ has no more than $n + 2(d - 1)k < n + 2dk = n + 4k(r - 1)^2$ break points. Since $s \in C^{r-2}[0, 1]$, all these break points are single knots. \[\square\]

**Remark.** The idea used in this proof can lead to a shorter proof of Theorem 1. But it will give a messier knot sequence and larger constant $C$, and a spline thus produced will not be as easy to handle as the one produced by our algorithm.

**REFERENCES**


