Unicity in Approximation of a Function and its Derivatives

By Lee Johnson

For \( f \) continuous and real on \([0, 1]\), let \( ||f|| = \max |f(x)|, \ x \in [0, 1] \). In this journal, Moursund [3] proved

**Theorem 1.** Let \( f \) be twice differentiable on \([0, 1]\). Among all polynomials \( h(x) \) of degree \( n \) or less, let \( p(x) \) be the one that minimizes: \( \max \{||h - f||, ||h' - f'||\} \). If \( q(x) \) is another such minimizing polynomial, then \( q' = p' \).

Let \( f^i \) denote the \( i \)th derivative of \( f \). Moursund’s result can be extended to:

**Theorem 2.** Let \( f \) be \((k + 1)\)-times differentiable on \([0, 1]\). Among all polynomials \( h(x) \) of degree \( n \) or less, let \( p(x) \) be the one that minimizes:

\[
\max \{||h - f||, ||h^1 - f^1||, \ldots, ||h^k - f^k||\}.
\]

If \( q(x) \) is another such minimizing polynomial, then \( q^k = p^k \).

We need some preliminary results before establishing Theorem 2. Let \( M(h) = \max \{||h||, \ldots, ||h^k||\} \). The functional \( M \) is a norm on the set \( S \) of functions that are \((k + 1)\)-times differentiable on \([0, 1]\).

Let \( Q \) denote the set of polynomials of degree \( n \) or less. Call \( p_0 \in Q \) a best approximation to \( f \in S \) if \( M(p_0 - f) \leq M(q - f) \), for all \( q \in Q \). It can be shown [1] that the set of best approximations is convex and nonempty.

Call \( x \in [0, 1] \) an extreme point of \( p - f \) if for some \( i \), \( 0 \leq i \leq k \), \( |p^i(x) - f^i(x)| = ||p^i - f^i|| = M(p - f) \). Denote the set of extreme points of \( p - f \) by \( E(p, f) \). Standard arguments quickly show [2] that \( p \) is a best approximation to \( f \) if and only if \( p \) is a best approximation to \( f \) on \( E(p, f) \).

**Proof of Theorem 2.** Let \( p \) and \( q \) be two best approximations to \( f \); and suppose \( p^k \neq q^k \). Let \( c = tp + (1 - t)q, \ t \in (0, 1) \); then \( c \) is also a best approximation to \( f \). Using \( p^k \neq q^k \), we will construct an approximation to \( f \) that is better than \( c \) on \( E(c, f) \), giving a contradiction. Let \( a_i = j \) if there are \( j \) points \( x \) in \((0, 1)\) such that \( |c^i(x) - f^i(x)| = ||c^i - f^i|| = M(c - f) \).

Let \( b_i = 0, 1, 2 \) according as none, one or both of \( z = 0, z = 1 \) are such that \( |c^i(z) - f^i(z)| = M(c - f) \). In particular, \( a_i = b_i = 0 \) if \( ||c^i - f^i|| < M(c - f) \).

If \( x_0 \) is among the \( a_i \) extreme points of \( c^i - f^i \), then

(1) \( x_0 \) is not among the \( a_{i+1} \) extreme points of \( c^{i+1} - f^{i+1} \),

(2) \( p^i(x_0) - f^i(x_0) = q^i(x_0) - f^i(x_0) = \pm M(c - f) \),

(3) \( p^{i+1}(x_0) - f^{i+1}(x_0) = q^{i+1}(x_0) - f^{i+1}(x_0) = 0 \).

From (2) and (3), \( p^i(x) - q^i(x) \) has at least \( 2a_i + b_i \) zeroes. We will show that \( p^i - q^i \) has at least \( (b_0 + \cdots + b_i) + 2(a_0 + \cdots + a_i) - i \) zeroes.

**Lemma.** Let \( h(x) \) be a polynomial with \( r \) single zeroes, \( s \) double zeroes and \( t \) triple zeroes. Let \( h'(x) \) have \( u \) double zeroes—none of which are among the \( t \) triple zeroes of \( h(x) \). Then \( h'(x) \) has at least \( r + 2s + 3t + 2u - 1 \) zeroes.

**Proof.** Let \( r + s + t = v \), and label the zeroes of \( h(x) \) as \( x_1, \ldots, x_r \); \( x_i < x_{i+1} \).

Received November 16, 1967.
In \((x_i, x_{i+1})\) there is a zero of \(h'(x)\); furthermore, this zero must be of odd multiplicity. Also none of the \(u\) double zeroes of \(h'(x)\) are counted among the \(v\) distinct zeroes of \(h(x)\). Counting the zeroes of \(h'(x)\) we obtain

(a) \(s + 2t\); from the multiple zeroes of \(h(x)\),
(b) \(v - 1\); the zeroes of \(h'(x)\) in \((x_i, x_{i+1})\),
(c) \(2u\); as noted, the \(v - 1\) zeroes in (b) are of odd multiplicity. If one of the \(u\) double zeroes of \(h'(x)\) is included in (b), this zero must have been of multiplicity \(3\) or more.

Adding (a), (b) and (c), establishes the lemma.

Using (1), (2) and (3) from above; and applying the lemma repeatedly to the derivatives of \(p(x) - q(x)\), we obtain

\[
(4) \quad p'(x) - q'(x) \text{ has at least } (b_0 + \cdots + b_i) + 2(a_0 + \cdots + a_i) - i \text{ zeroes.}
\]

As \(p^k - q^k \neq 0\), it must be that \(n - k \geq (b_0 + \cdots + b_k) + 2(a_0 + \cdots + a_k) - k\).

The same argument, starting with \(p^i(x) - q^i(x)\), gives that \(p^k - q^k\) has \((b_j + \cdots + b_k) + 2(a_j + \cdots + a_k) - (k - j)\) zeroes. Thus, \(p^k \neq q^k\) means that

\[
(5) \quad (b_j + \cdots + b_k) + 2(a_j + \cdots + a_k) \leq n - j, \quad 0 \leq j \leq k.
\]

We will use (5) to construct a polynomial \(r(x)\) such that \(r'(y) - f'(y) = 0\) if \(y\) is one of the \(a_i + b_i\) extreme points of \(c_i - f_i\). Select \(s\) points in \((0, 1)\), distinct from the \(a_0 + b_0\) extreme points of \(c - f\); where \(s + (b_0 + \cdots + b_k) + 2(a_0 + \cdots + a_k) - n + 1\). Note that from (5), \(s \geq 1\).

Let \(D(x) = (1, x, x^2, \ldots, x^n)\), \(D^i(x) = (0, 1, 2x, \ldots, nx^{n-1})\), \(D^2(x) = (0, 0, 2, \ldots, n(n - 1)x^{n-2})\). Define \(D^i\) similarly, \(i = 3, \ldots, k\). We will form an \((n + 1) \times (n + 1)\) "Vandermonde-like" matrix \(A\), as follows. For each of the \(s\) points \(y_1, \ldots, y_s\) chosen in \((0, 1)\), let \(A\) have a row of the form \(D(y_i)\). For each of the \(a_i\) extreme points \(w\) of \(c - f\), let \(A\) have two rows of the form

\[
\begin{align*}
D(w) \\
D'(w)
\end{align*}
\]

For each of the \(b_i\) "end-point" extreme points \(z\) of \(c - f\), let \(A\) have a row of the form \(D(z)\).

Generally, for each of the \(a_i\) extreme points \(w\) of \(c^i - f^i\), let \(A\) have 2 rows of the form

\[
\begin{align*}
D^i(w) \\
D^{i+1}(w)
\end{align*}
\]

For each of the \(b_i\) "end-point" extreme points of \(c^i - f^i\), let \(A\) have a row of the form \(D^i(z)\).

We now show that \(A\) is nonsingular. Suppose \(Ad^T\) is the zero vector; where \(d = (d_0, d_1, \ldots, d_n)\). Form \(h(x) = d_0x^n + \cdots + d_1x + d_0\). Clearly, \(h^i(x)\) has \(a_i\) double zeroes and \(b_i\) single zeroes; \(0 \leq i \leq k\). Applying the lemma, \(h^k(x)\) has \(s + (b_0 + \cdots + b_k) + 2(a_0 + \cdots + a_k) - k = n + 1 - k\) zeroes. As \(h^k(x)\) has degree \(n - k\) or less, \(h^k = 0\).

Using (5), and \(s + (b_0 + \cdots + b_k) + 2(a_0 + \cdots + a_k) = n + 1\), we have

\[s + (b_0 + \cdots + b_j) + 2(a_0 + \cdots + a_j) \geq j + 2; \quad j = 0, 1, \ldots, k - 1.\]

Hence by (4), \(h^j(x)\) has at least 2 zeroes, \(j = 0, 1, \ldots, k - 1\). As \(h^k = 0\), this
shows that \( h = 0 \), or \( d_i = 0 \), \( i = 0, 1, \ldots, k \). Thus, \( A \) is nonsingular.

As \( A \) is nonsingular, we can fit \( f(x), f'(x), \ldots, f^k(x) \) exactly on the \( (b_0 + \cdots + b_k) + 2(a_0 + \cdots + a_k) \leq n \) extreme points of \( c - f \). That is, we can find \( r(x) \) of degree \( n \) or less, so that if \( |c'(x') - f'(x')| = ||c^i - f^i|| = M(c - f) \), then \( r'(x') - f'(x') = 0 \). It may well be, even though \( r'(x') - f'(x') = 0 \), that \( |r^i(x') - f^i(x')| \geq M(c - f) \) for some \( j, j \neq i \). If this is the case, \( x' \) must not have been one of the \( a_j + b_j \) extreme points of \( c^i - f^i \). If \( |c'(x') - f'(x')| < M(c - f) \), there is some \( t \in (0, 1) \) such that

\[
|t(r^i(x') - f^i(x')) + (1 - t)(c^i(x') - f^i(x'))| < M(c - f).
\]

As \( E(c, f) \) was supposed to be a finite set, we can use the above remark to choose some \( t \in (0, 1) \) such that

\[
|t(r^i(x) - f^i(x)) + (1 - t)(c^i(x) - f^i(x))| < M(c - f), \quad \text{for all } x \in E(c, f), \quad i = 0, 1, \ldots, k.
\]

This gives \( tr + (1 - t)c \) a better approximation to \( f \) on \( E(c, f) \) than is \( c \). Thus, \( c \) could not have been a best approximation.

The proof above, except for cumbersome notational modifications, clearly establishes the more general

**Theorem 3.** Let \( i, j, \ldots, k \) be any finite sequence of nonnegative integers, \( i < j < \cdots < k \). Let \( f(x) \) be \( (k + 1) \)-times differentiable on \( [a, b] \). Among all polynomials \( h(x) \) of degree \( n \) or less, let \( p(x) \) be one that minimizes:

\[
\max \{ ||h^i - f^i||, ||h^j - f^j||, \ldots, ||h^k - f^k|| \}.
\]

If \( q(x) \) is another such minimizing polynomial, then \( q^k = p^k \).