JOURNAL OF APPROXIMATION THEORY **64**, 17–24 (1991)

Best Interpolation with Polynomials Lacking the Linear Term

THEODORE KILGORE

Department of Algebra, Combinatorics and Analysis, Auburn University, Alabama 36849

Communicated by R. Bojanic

Received April 10, 1986; revised February 27, 1990

Interpolation of minimal norm into the space of polynomials spanned by the monomials $1, t^2, ..., t^{n+1}$ on an interval $[a, b]$, with $0 < a < b$, is shown here to be characterized by the equioscillation properties (Bernstein–Erdős conditions) which characterize minimal norm interpolation into several other spaces, including in particular the classical case of the space spanned by the consecutive monomials $1, t, ..., t^n$ on any underlying interval. A natural conjecture is that the Bernstein and Erdős conditions in fact characterize minimal norm interpolation in practically any space spanned by an extended Chebyshev system (Markov system). Thus far, this conjecture eludes proof. It is hoped that the techniques employed here, in a concrete case, may be of help in this more general problem.

1. INTRODUCTION

Spaces of incomplete or lacunary polynomials present an area of interesting complexity in the investigation of optimal interpolation. It has been conjectured [6] that optimal interpolation from $C[a, b]$ into such spaces is characterized by the conditions of Bernstein and Erdős, described below, which characterize optimal interpolation into several other types of spaces. Indeed, this has been shown true [6] for incomplete polynomials spanned by the monomials $1, t^{k+1}, ..., t^{k+n}$ when the left endpoint, a, of the interval of interpolation is 0. A problem of much greater complexity arises if, with a natural desire for generality, one allows $a > 0$ as well. We state the following result here.

2. STATEMENT OF RESULT

THEOREM. *For $n \geqslant 2$, let* **Y** *be the space of polynomials spanned by* $1, t^2, ..., t^{n+1}$. *Optimal interpolation with this space on an interval* $[a, b]$,

0021-9045/91 $3.00

with $0<a<b$, *and with nodes* $t_0, \ldots, t_n$, *satisfying* $a=t_0<\cdots<t_n=b$, *is characterized by the conditions of Bernstein and Erdős, and the nodes* $t_0, \ldots, t_n$ *are uniquely determined.*

Remark. The following proof also shows slightly more. It shows that the norm of optimal interpolation with this space, for any given n, must exceed the norm of optimal interpolation with polynomials of degree $n-1$. As discussed in [8], this result follows directly from the nature of the proof given below.

3. Method of Proof

We first introduce necessary notation and describe the Bernstein and Erdős conditions mentioned above. An outline of the proof of the theorem concludes this section.

The problem solved here is one of a class of problems with some common features, and its proof may best be motivated in the general context. We begin, therefore, with some general observations, which parallel the development in [9], where the methodology applied here was laid out.

We let the space $\mathbf{Y}$ of dimension $n+1$ in $C[a, b]$ be spanned by a complete extended Tchebycheff system. Then given nodes $t_0, \ldots, t_n$ in $[a, b]$ such that $a=t_0<\cdots<t_n=b$, there exists a basis $\{y_0, \ldots, y_n\}$ for $\mathbf{Y}$ such that

$$y_i(t_j)=\delta_{ij} \qquad \text{(Kronecker delta)}.$$

An interpolating projection $P\colon C[a, b] \to \mathbf{Y}$ may then be defined for $f \in C[a, b]$ by

$$Pf=\sum_{i=0}^{n} f(t_i)\, y_i.$$

It is seen that P is bounded, and

$$\|P\|=\left\|\sum_{i=0}^{n} |y_i|\right\|.$$

The Lebesgue function

$$\Lambda=\sum_{i=0}^{n} |y_i|$$

has the properties that

$$\Lambda(t)=1 \qquad \text{if } t \text{ is a node},$$

and, for $i \in \{1, ..., n\}$,

> Λ has a unique maximum T_i on the subinterval $[t_{i-1}, t_i]$ between two consecutive nodes (provided that $n > 1$), at which points Λ is differentiable, and $\Lambda'(T_i) = 0$.

Clearly, $\|P\| = \max\{\Lambda(T_1), ..., \Lambda(T_n)\}$, and $\|P\|$ depends upon the choice of nodes $t_1, ..., t_{n-1}$.

That the norm of P is minimal if

$$\Lambda(T_1) = \cdots = \Lambda(T_n) = C_Y,$$

for some unique value C_Y, the equality holding on a uniquely determined set of nodes, is a natural generalization of the Bernstein conjecture on Lagrange interpolation [1]. That, furthermore, if $\|P\| > C_Y$, at least one of the local maximum values of Λ is less than C_Y is an equally natural generalization of the Erdős conjecture on Lagrange interpolation [3], and these conditions combined are the conjecture of [6]. Our theorem states that both of these generalized conjectures are valid and characterize optimal interpolation for the space under consideration.

We define

$$\lambda_i = \Lambda(T_i) = \max \Lambda(t)_{[t_{i-1}, t_i]}, \qquad i \in \{1, ..., n\},$$

and denote by X_i the function in Y which agrees with Λ on $[t_{i-1}, t_i]$. The derivative of the function from $\mathbf{R}^{n-1}$ to $\mathbf{R}^n$ given by

$$(t_1, ..., t_{n-1}) \mapsto (\lambda_1, ..., \lambda_n)$$

exists and is given by a matrix

$$\left(\frac{\partial \lambda_i}{\partial t_j}\right)_{i=1}^{n}{}_{j=1}^{n-1}. \tag{1}$$

We denote by J_p the determinant of the square matrix derived by removing the pth row, for each $p \in \{1, ..., n\}$.

To establish the generalized Bernstein and Erdős conjectures of [6] as valid characterizations of optimal interpolation into $\mathbf{Y}$, it suffices to show [4, 2, 5, 10] that

(i) $J_p \neq 0$ for all possible choices of the nodes and for

$$p \in \{1, ..., n\},$$

and

(ii) J_p alternates in sign.

The equivalence

$$\partial\lambda_i/\partial t_j = -y_j(T_i)\, X_i'(t_j) \tag{2}$$

facilitates our work. Using the methods described in [9], we reduce this matrix by column and row cancellations to an equivalent matrix

$$(q_i(t_j))_{i=1}^{n}{}_{j=1}^{n-1}, \tag{3}$$

reducing (i) and (ii) to a question of whether the set of functions $\{q_1, ..., q_n\} \backslash \{q_p\}$, $p \in \{1, ..., n\}$, admits a non-trivial linear combination which is zero on the points $t_1, ..., t_{n-1}$. The proof is then completed by answering this question.

Proof of the Theorem. As discussed above, it is necessary to construct the matrix (1) of partial derivatives and to establish the determinant properties (i) and (ii). We begin by obtaining explicit expressions for the fundamental functions. We may write explicitly for $i \in \{0, ..., n\}$

$$y_i(t) = \left(\prod_{\substack{j=0 \\ j \neq i}}^{n} \frac{t - t_j}{t_i - t_j}\right)\left(\frac{f_i(t)}{f_i(t_i)}\right),$$

where

$$f_i(t_i) = \sum_{j=0}^{n} \prod_{\substack{l=0 \\ l \neq j}} t_l, \tag{4}$$

and where $f_i(t)$ is obtained by replacing t_i with t in the above formula. We note that $f_i(t_i) = f_j(t_j)$ for $i, j \in \{0, ..., n\}$.

The functions $f_l(t)$ are linear, symmetric, and positive for positive $t_0, ..., t_n$, and t. For $j \neq l$, we write $f_l(t; t_j)$ to denote that t_j is the independent variable, the others being held constant. The important identity

$$f_i(t; t_j)|_{t_j = s} = f_j(t; t_i)|_{t_i = s} \tag{5}$$

implies that $f_l(t; t_j)$ may be viewed as a linear function with a negative root which would move to the left as t increases on the interval $(0, \infty)$, as may be seen from the following computation, which cuts through the notational complexity by relabelling the points. We assume without loss of generality that $t_0, ..., t_{n-1}$ are positive and that t_n is such that

$$\sum_{l=0}^{n} \prod_{\substack{m \neq l \\ m=0}}^{n} t_m = 0.$$

Then

$$t_n = \left(-\prod_{m=0}^{n-1} t_m \right) \left(\sum_{l=0}^{n-1} \prod_{\substack{m \neq l \\ m=0}}^{n-1} t_m \right)^{-1}$$

and, after application of the quotient rule, the numerator of $\partial t_n / \partial t_j$, for any $j \in \{0, ..., n-1\}$, is equal to

$$-\sum_{\substack{m=0 \\ m \neq j}}^{n-1} t_m^2,$$

which is negative.

Using the equivalence (2) to rewrite the matrix (1), we may carry out the matrix manipulations described in [9], reducing the matrix (1) in this context to a matrix of form (3), in which we may define the entries $q_i(t_j)$, for $i \in \{1, ..., n\}$, $j \in \{1, ..., n-1\}$ by

$$q_i(t_j) = \left(\prod_{\substack{k=1 \\ k \neq i}}^{n} \prod_{\substack{l=1 \\ l \neq j}}^{n-1} f_l(T_k; t_j) \right) \frac{X_i'(t_j)}{t_j(t_j - T_i)}, \tag{6}$$

in which by (5) we may regard $q_1, ..., q_n$ as polynomials evaluated at successive points $t_1, ..., t_{n-1}$ across the rows of the matrix. Only the representations of these functions changes from column to column.

It is now possible to ascertain conditions (i) and (ii) by looking at the locations of the roots of the polynomials $q_1, ..., q_n$, in a manner similar to that used in [7]. We note first of all that each of the polynomials $X_1', ..., X_n'$ has a root at zero, which may be cancelled, as indicated in (6), without affecting the nonsingularity properties on the nodes. All other roots of these polynomials are real, and on the interval $[T_1, T_n]$ their roots strictly interlace in a cyclic pattern. Moreover, the polynomials X_1' and X_n' each have $n-1$ roots on this interval. The other polynomials $X_2', ..., X_n'$ each have exactly $n-2$ roots on the interval, and perhaps another root in some location outside of the interval. Under certain circumstances which will be described below, the location of this root may cause a problem. A method for dealing with this problem will be introduced there. We adopt, for $i \in \{1, ..., n\}$, the more compact notation

$$Q_i(t) = t^{-1}(t - T_i)^{-1} X_1'(t),$$

emphasizing again that $Q_i(t)$ is a polynomial.

We move now to consideration of the factors $f_l(T_k)$, for $l \in \{1, ..., n-1\}$, $k \in \{1, ..., n\}$. For convenience, and with no loss of generality, we will choose to use the particular representation of the functions (6) which occurs in the first column.

For each $l \in \{2, ..., n-1\}$, the functions

$$f_l(T_k; t_1), \qquad k \in \{1, ..., n\}$$

in the variable t_1 have roots $s_1, l, ..., s_n, l$ such that

$$s_{n,l} < s_{n-1,l} < \cdots < s_{1,l} < 0, \tag{7}$$

and the factor

$$\prod_{\substack{k=1 \\ k \neq i}}^{n} f_l(T_k; t_1)$$

which appears in the i, 1st entry has roots on the set

$$\{s_{1,l}, ..., s_{n,l}\} \setminus \{s_{i,l}\}.$$

We now adopt as a simplification of our problem a standard representation of the polynomials $q_1, ..., q_n$ by writing

$$q_i(t) = c_i \prod_{l=2}^{n-1} \prod_{\substack{k=1 \\ k \neq i}}^{n} (t - s_{k,l})\, Q_i(t) \tag{8}$$

in which $c_1, ..., c_n$ are whatever non-zero constants are appropriate.

We now begin the demonstration that (i) is true by successive reduction of the degree of the polynomials. The method will involve adding suitable polynomials to $q_1, ..., q_n$, causing the roots $s_{k,l}$ to coalesce in such a way that the corresponding factors can be cancelled from the columns, resulting at last in the reduction to an equivalent matrix the numerical value of whose entries is unchanged, but whose entries are now represented by polynomials of degree not exceeding $n-1$, evaluated at the original points of evaluation. Matrices of this form are known to have the nonsingularity properties (i) and (ii). The demonstration below is quite similar to that used in [7]. It will be necessary to discuss two cases in order to complete the argument.

We will show, for a fixed but arbitrary index $l \in \{2, ..., n\}$, that the roots $s_{1,l}, ..., s_{n,l}$ may be moved successively to the location of $s_{n,l}$ and cancelled. To facilitate the presentation, we rewrite (8) for $i \in \{1, ..., n\}$ as

$$q_i(t) = L_i(t)\, R_i(t) \tag{9}$$

in which R_i is the polynomial defined by

$$R_i(t) = \prod_{\substack{k=1 \\ k \neq i}}^{n} (t - s_{k,l})\, Q_i(t)$$

and by L_i we denote the polynomial whose roots are the remainder of the roots of q_i. It is now necessary to distinguish two cases. We recall that the polynomials $Q_2, ..., Q_{n-1}$ may each possess a single root which lies outside of the interval $[T_1, T_n]$. We assume as the first case that none of these roots lie on the interval $[s_{n,l}, T_1)$. If any of these roots are so situated, we must move to the second case, in which we describe a method by which they can be moved away from that interval.

We observe that for $j \in \{1, ..., n-1\}$,

$$q_i(t_j) = L_i(t_j)(P_i(t_j) + R_i(t_j)) \tag{10}$$

whenever P_i is a polynomial which has roots at the points $t_1, ..., t_{n-1}$. Thus, for $i \in \{1, ..., n\}$ we choose P_i to be the polynomial of minimal degree which has roots at $t_1, ..., t_{n-1}$ and at each point in the set $\{s_{1,l}, ..., s_{n-2,l}\} \setminus \{s_{i,l}\}$, and we assign to P_i the value

$$P_i(s_{n,l}) = -R_i(s_{n,l}).$$

We note that P_i is identically zero if $R_i(s_{n,l}) = 0$. Otherwise, the result of inserting this particular P_i in (10) is a polynomial whose degree does not exceed that of the original q_i for $i=1$ or $i=n$. For $i \in \{2, ..., n-1\}$, the degree of the new polynomial may indeed exceed the degree of the old polynomial by one, but under no circumstances does it exceed the degree of q_1 or q_n. Moreover, the sign of P_i agrees with the sign of R_i on the set $\{T_1, ..., T_n\} \setminus \{T_i\}$, implying that the sign of q_i cannot change at these points. The sign cannot change at T_i either because the number of roots of $P_i + R_i$ would exceed its degree. Thus, the degree of the polynomials $q_1, ..., q_n$ may be decreased by cancelling the factor $(t_j - s_{n,l})$ from the jth column of the matrix for $j \in \{1, ..., n\}$. The argument may now be repeated until all of the roots $s_{1,l}, ..., s_{n,l}$ have been removed from the polynomials in the matrix, leaving a set of polynomials of degree $n-1$ or less which preserve their original signs at the points $T_1, ..., T_n$. As stated, this condition suffices to establish determinant property (i).

We now turn our consideration to the second case, in which at least one of the polynomials $Q_2, ..., Q_{n-1}$ possesses a root on the interval $[s_{n,l}, T_1)$. For each $i \in \{1, ..., n\}$ such that Q_i has no root on this interval, we add to Q_i the polynomial

$$P_i(t) = b_i(t-t_1)(t-t_2) \cdots (t-t_{n-1}),$$

where b_i is of such sign as to make the expression P_i negative at the point T_1 and is of sufficiently small magnitude that the crucial sign properties on the points $T_1, ..., T_n$ are not violated by $Q_i + P_i$, which we immediately relabel as Q_i. The effect of this small perturbation of the system of polynomials is to provide for each of them a root to the left of T_1 and, for $i \in \{2, ..., n-1\}$, another to the right of T_n, and, of course, to increase the degree of those polynomials so treated to $n-1$. The constant b_1 should have been chosen with more care than the others, as the new root lying to the left of T_1 of the new polynomial Q_1 should be the leftmost of all of the roots of all of the new polynomials $Q_1, ..., Q_n$ and should furthermore be to the left of the point $s_{n,l}$. This, however, can be done; if the other coefficients are chosen first, we simply have a second condition which requires an upper bound on $|b_1|$. At this juncture, the roots of $Q_2, ..., Q_n$ which lie to the left of T_1 can be moved to a common location with the leftmost root of Q_1 and cancelled from the matrix by arguments essentially identical to those used in the first case, discussed above. After this cancellation, the second case has been reduced to the first.

This argument concludes the proof of the properties (i) and (ii), which have been shown to imply the theorem.

4. Concluding Remark

It is hoped that this paper will serve as an opening to more general problems of optimization of interpolation with incomplete or lacunary polynomials.

References

1. S. Bernstein, Sur la limitation des valeurs d'un polynome $P_n(x)$ de degré n sur tout un segment par ses valeurs en $n+1$ points du segment, *Izv. Akad. Nauk SSSR* **7** (1931), 1025–1050.
2. C. De Boor and A. Pinkus, Proof of the conjectures of Bernstein and Erdős concerning polynomial interpolation, *J. Approx. Theory* **24** (1978), 289–303.
3. P. Erdős, Some remarks on polynomials, *Bull. Amer. Math. Soc.* **53** (1947), 1169–1176.
4. T. A. Kilgore, Optimization of the Lagrange interpolation operator, *Bull. Amer. Math. Soc.* **83** (1977), 1069–1071.
5. T. A. Kilgore, A characterization of the Lagrange interpolating projection with minimal Tchebycheff norm, *J. Approx. Theory* **24** (1978), 273–288.
6. T. A. Kilgore, Optimal interpolation with incomplete polynomials, *J. Approx. Theory* **41** (1984), 279–290.
7. T. A. Kilgore, Optimal interpolation with polynomial having fixed roots, *J. Approx. Theory* **49** (1987), 378–389.
8. T. A. Kilgore, A lower bound on the norm of interpolation with an extended Tchebycheff system, *J. Approx. Theory* **47** (1986), 240–245.
9. T. A. Kilgore, Optimal interpolation in a general setting—A beginning, *J. Approx. Theory* **55** (1988), 289–295.
10. H. Loeb, A differential equations approach to the Bernstein problem, *in* "Numerical Methods of Approximation Theory" (L. Collatz, G. Meinardus, and H. Werner, Eds.), Birkhaüser, Basel, 1980.