CONSTRUCTIVE THEORY OF FUNCTIONS'87

SOFIA, 1988

SOME RECENT RESULTS IN OPTIMAL INTERPOLATION

Theodore Kilgore

Introduction: The Bernstein [1] and Erdos [7] conjectures dealt with the nodes which produce Lagrange interpolation of minimal norm. These conjectures are apparently applicable to interpolation into almost any reasonable space. Indeed, the conditions laid down by Bernstein and Erdos in their conjectures do characterize optimal interpolation into various spaces of polynomials, rational functions, and trigonometric functions, as will be described in our further discussion. Even a step has been made toward showing that the conjectures apply with validity to range spaces spanned by extended complete Tchebycheff systems, (spanned by functions $u_0, \ldots, u_n$ such that the Wronskian of $u_0, \ldots, u_k$ is nonvanishing in $[a, b]$ for $k = 0, \ldots, n$). However, the remainder of the proof under such generality remains elusive. We wish to review here the methods used in some extensions of the Bernstein and Erdos conjectures and to highlight some of the problems encountered in these and perhaps future attempts to generalize their application. A discussion of the problems encountered in dealing with certain spaces spanned by polynomials, especially lacunary or incomplete polynomials, should suffice for purposes of such a survey. Before doing so, however, we will give some attention to notation and terminology and also restate the conjectures of Bernstein and Erdos.

Terminology and Notation: Although we wish to deal with polynomial spaces, certain basic definitions are most easily given in a general context. Therefore, let Y be an $n+1$ dimensional subspace of $C[a, b]$ spanned by an extended complete Tchebycheff system. Let points (nodes) $t_0, \ldots, t_n$ be given such that $a = t_0 < t_1 < \ldots < t_n = b$, and let $y_0, \ldots, y_n$ be a fundamental basis for Y, such that $y_i(t_j) = \delta_{ij}$ (Kronecker delta). Then the linear projection P: $C[a, b] \to Y$ defined for $f \in C[a, b]$ by $Pf = \sum_{i=0}^{n} f(t_i)\, y_i$ is the interpolating projection based on the nodes $t_0, \ldots, t_n$, and it may be seen that

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

(we are using the sup norm). The function defined by the sum on the right is called the Lebesgue function of P and is customarily denoted by $\Lambda(t)$. Its value is 1 at each node, and it has a maximum T_i with value $\Lambda(T_i) = \lambda_i$ on each subinterval $(t_i - t_{i-1})$ of $[a, b]$, $i = 1, 2, \ldots, n$.

The conjecture of Bernstein and Erdos on Lagrange interpolation:

A clarifying modification of the conjecture of Bernstein [1] is that $\|P\|$ is minimal when $\lambda_1 = \ldots = \lambda_n$ (a phenomenon sometimes called "equioscillation"), to which we can add that this occurs at a unique set of nodes. Erdos [7] conjectured that moreover there exists a value C associated with Y, such that, if for some i, $\lambda_i > C$, then there is a j such that $\lambda_j < C$.

As stated, these two conjectures applied specifically to Lagrange interpolation. The methods by which they were proved were announced in Kilgore [8] and employed in Kilgore [9] to prove the Bernstein conjecture and to show that the norm of optimal Lagrange interpolation increases strictly as n increases. De Boor and Pinkus employed the method of proof outlined in [8] to give proofs of the Bernstein and Erdos conjectures, their article [2] appearing simultaneously with [9]. In [2], de Boor and Pinkus also adapted slightly the arguments used, establishing a similar characterization of optimal trigonometric interpolation of continuous 2π - periodic functions. This rather natural extension of the original problem had also been investigated by Kilgore, who did not publish his efforts after the appearance of [2]. Since the appearance of these three papers, the criteria conjectured by Bernstein and Erdos have been shown to characterize optimal interpolation for several other choices of the range space (see References)..

Historical comments:

The solutions of the Bernstein and Erdos conjectures have begun to appear in the literature of approximation theory. One book [16, p. 318] contains a brief mention of the problem, in which it is stated "The solution ... was conjectured by Bernstein in 1931, but the conjecture was not proved until 1977, by de Boor and Pinkus and by Kilgore independently." A second book [4, Appendix] contains a more exhaustive treatment of the solution, and states the following historical account:

> In 1976, Kilgore and Cheney found that, for each $n > 1$ there is a set of knots with an equioscillating Lebesgue function. Finally, in 1978 Kilgore showed that the equioscillation property is a necessary condition. On the basis of this local result and by applying topological arguments, de Boor and Pinkus (1978) completed the proof of the conjectures. Notes on the interesting history ... are found in the cited literature.

This account is incomplete, and, in the interests of historical completeness, I would like to make the following observations, which can be checked by a perusal of the "cited literature."

1) This account omits mention of Kilgore [8], which appeared in September, 1977 and was a source for both [9] and [2].

2) Kilgore [9] was received for publication on March 12, 1977 and appeared in December, 1978.

3) De Boor & Pinkus [2] was received for publication on April 1, 1977, also appearing December 1978.

4) As noted by de Boor and Pinkus [2] and later by de Boor [3], Kilgore [8] contained a proof of the "necessity" referred to, and Kilgore [9] contained a complete proof of the Bernstein conjecture. The casual reader of [4] might be left with a contrary impression.

Methods of proof of the Bernstein and Erdos conjectures, applied to various range spaces:

As stated in the introduction, our object here is to describe applications of the Bernstein and Erdos characterization of optimal interpolation to various choices of Y showing an evolution in the arguments employed. In a very wide context, which certainly includes polynomials or other analytic functions, the location of the points $T_1, \ldots, T_n$ varies in continuously differentiable fashion if the nodes $t_1, \ldots, t_{n-1}$ are moved, and we have the formula

$$\partial\lambda_i/\partial t_j = -y_j(T_i)\, X_i'(t_j) \tag{1}$$

for $i=1, \ldots, n$ and $j=1, \ldots, n-1$. The function X_i is defined in this connection to be that function in Y given by

$$X_i(t) = \sigma_{i0}\, y_0(t) + \sigma_{i1}\, y_1(t) + \ldots + \sigma_{in}\, y_n(t)$$

for t in $[a, b]$, where $\sigma_{ij} = \operatorname{sgn} y_j(T_i)$; we note that $X_i(T_i) = \lambda_i$; more generally $X_i(t) = \Lambda(t)$ for $t_{i-1} \leq t \leq t_i$; and $X_i'(T_i) = 0$. Formula (1) derives from [15]. Its proof is too long to be reproduced here, but the reader who desires to reconstruct it should note merely that, by the chain rule for partial derivatives,

$$\partial\lambda_i/\partial t_j = \frac{\partial(\text{coefficients of } X_i)}{\partial t_j} + X_i'(T_i)\, \partial T_i/\partial t_j,$$

and the second term vanishes because $X_i'(T_i) = 0$. The derivation of (1) may now be completed by noting that, if $y_0, \ldots, y_n$ interpolate on the nodes

$$t_0, \ldots, t_{j-1}, t_j, t_{j+1}, \ldots, t_n$$

and $z_0, \ldots, z_n$ interpolate on $t_0, \ldots, t_{j-1}, s, t_{j+1}, \ldots, t_n$,

and $P_1(t) = A_0\, y_0(t) + \ldots + A_n\, y_n(t)$,

$P_2(t) = A_0\, z_0(t) + \ldots + A_n\, z_n(t)$

then $P_1(t) - P_2(t)$ is a multiple of $y_j(t)$ (or $z_j(t)$).

Now, the Bernstein and Erdos conditions for optimal interpolation will follow by standard topological arguments if the matrix of partial derivatives

$$\left(\partial\lambda_i/\partial t_j\right)_{j=1\ i=1}^{n-1\ n} = \left(-y_j(T_i)\, X_i'(t_j)\right)_{j=1\ i=1}^{n-1\ n} \tag{2}$$

satisfies (i) and (ii):

(i) for $k = 1, \ldots, n$, $J_k = \det \left(\partial\lambda_i / \partial t_j \right)_{j=1 \; i=1, \, i \neq k}^{n-1 \; n}$ is not zero.

(ii) for $k \varepsilon 1, \ldots, n$, $(-1)^k J_1 J_k < 0$.

Condition (ii) may be seen in the context of the interpolation problem to follow from (i). Our proof therefore hinges on the establishment of (i).

<u>The case of Lagrange interpolation:</u>

The portion of the proof which has been presented up to this point is obviously very general in nature. A divergence now appears between different choices of the range space Y in the establishing of the crucial condition (i). We will begin by recapitulating what happens in the classical case of Lagrange interpolation.

If Y is the space of polynomials of degree n or less, we have the formula

$$y_j(t) = \prod_{\substack{\ell=0 \\ \ell \neq j}}^{n} (t - t_\ell)(t_j - t_\ell)^{-1} \tag{3}$$

Using this formula to express the $y_j(T_i)$ portion of the $i\underline{j}\underline{th}$ entry of (2), we may cancel from the $j\underline{th}$ row of (2), for $j=1, \ldots, n-1$, the denominator of y_j. We may then, for $i=1, \ldots, n$, divide each entry in the $i\underline{th}$ column by $(T_i - t_0) \ldots (T_i - t_n)$. The resulting matrix, equivalent with (2), is

$$\left(\frac{X_i'(t_j)}{t_j - T_i} \right)_{j=1 \; i=1}^{n-1 \; n} \tag{4}$$

and we note that

$$q_i(t) := \frac{X_i'(t)}{t - T_i} \tag{4a}$$

is a polynomial of degree $n-2$ or less. Thus, condition (i) follows immediately if $\{q_1, \ldots, q_n\} - \{q_k\}$ is linearly independent for $k=1, \ldots, n$. These functions $q_1, \ldots, q_n$ in turn satisfy certain sign properties on the points $T_1, \ldots, T_n$, which properties do not depend on the particular choice of nodes:

(a) For all applicable i and ℓ, the function q_i must change sign (exactly once) on the interval $[T_{\ell-1}, T_\ell]$, except that

(b) The function q_i does not change sign (and is not zero) on $[T_{i-1}, T_i]$ $[T_i, T_{i+1}]$, and

(c) $q_i(T_\ell) \neq 0$ for all i and for all ℓ.

The proof of the Bernstein and Erdos conjectures for Lagrange interpolation is then completed by proving the following general statement:

<u>Proposition 1</u>: Let $q_1, \ldots, q_n$ be polynomials of degree $n-2$ or less, satisfying sign conditions (a) - (c) on an ordered set of points $T_1, \ldots, T_n$. Then for any $k \varepsilon \{1, \ldots, n\}$, the set $\{q_1, \ldots, q_n\} - \{q_k\}$ is linearly independent.

Proof: For details, the interested reader may consult [10]. We provide an outline.

First of all, we normalize the functions $q_1, \ldots, q_n$ by assuming them positive at T_1. We then consider a linear combination equal to zero:

$$Q = a_1 q_1 + \dots + a_n q_n = 0$$

in which, for some k, $a_k = 0$. If $k = 1$, we renumber the system from the right, so we assume $k \neq 1$, and we also assume that $a_1 \geq 0$. We now partition the index set $\{1, \dots, n\}$ into two nonvoid, disjoint subsets

$$\mathcal{S} = \{j: a_j \geq 0 \text{ and } j \geq 2\}, \text{ and}$$
$$\mathcal{R} = \{1, \dots, n\} - \mathcal{S}.$$

and we define

$$R = \sum_{j \in \mathcal{R}} a_j q_j \quad \text{and} \quad S = \sum_{j \in \mathcal{S}} a_j q_j.$$

We have $Q = R + S$.

Now, the properties (a) - (c) imply that, at each of the points $T_2, \dots, T_n$, $(-1)^i q_1(T_i) > 0$, and moreover for $i, j \in \{2, \dots, n\}$

$$q_1(T_i)\, q_j(T_i) < 0 \quad \text{if } j \neq i$$

while $q_1(T_i)\, q_i(T_i) > 0$. From these facts, we derive that both $R = 0$ and $S = 0$ separately. But then $S(T_1) = 0$ implies that $a_j = 0$ for $j \in \mathcal{S}$, and $R(T_k) = 0$ implies that $a_j = 0$ for $j \in \mathcal{R}$, and the proof is completed.

A corollary of Proposition 1 was used also in Kilgore [9]. We state this corollary in the generalized form which appears in [10]. It will be used below in the proof of Proposition 2.

<u>Corollary</u>: Let $q_1, \dots, q_n$ be polynomials of degree $n-2$ or less satisfying sign conditions (a) - (c) on points $T_1, \dots, T_n$. Let $k, \ell \in \{1, \dots, n\}$, $k \neq \ell$. Then no linear combination of $\{q_1, \dots, q_n\} - \{q_k, q_\ell\}$ may have zeroes between points T_j, T_{j+1} in the same fashion as the zeroes of q_k or q_ℓ.

<u>More general problems — the matrix reduction problem</u>:

In the above proof, the matrix (2) may be reduced to the form (4), and thus the matrix of partial derivatives may be viewed as a matrix in which functions $q_1, \dots, q_n$ are evaluated at points $t_1, \dots, t_{n-1}$. It is not always so easy to perform such a reduction, although a general method for doing so has been laid down in [13], which is valid so long as the space Y under consideration is spanned by some extended Tchebycheff system. In such an eventuality, we may write for $i = 0, \dots, n$

$$y_i(t) = \frac{U_i(t)}{U_i(t_i)} \prod_{\substack{j \neq i \\ j=0}}^{n} (t - t_j)(t_i - t_j)^{-1} \tag{5}$$

where the expression $U_i(t)$ depends upon t and $t_0, \dots, t_n$, excepting t_i, and is a symmetric function in the indicated variables. Futhermore, we have

$$U_i(t) \Big|_{t_j = s} = U_j(t) \Big|_{t_i = s}. \tag{6}$$

This identity enables us to reduce the matrix (2) to a matrix of evaluation as follows:

Carrying out the same cancellations as previously described, we do not end with (4), but rather with

$$\left(U_j(T_i)\ \frac{X_i'(t_j)}{t_j - T_i}\right)_{j=1\ i=1}^{n-1\ n} \tag{7}$$

We now divide the ith column by the product $U_1(t_i) \ldots U_{n-1}(T_i)$, ending with

$$\left(\prod_{\substack{\ell=1 \\ \ell \neq j}}^{n-1} (U_\ell(T_i))^{-1}\ \frac{X_i'(t_j)}{t_j - T_i}\right)_{j=1\ i=1}^{n-1\ n} \tag{8}$$

which is a matrix of evaluation, by (6). In the case that the space Y is spanned by polynomials, the functions $U_1, \ldots, U_{n-1}$ are also polynomials, and (8) consists of an evaluation matrix of rational functions. However, the entries of (8) can be replaced by polynomials, using appropriate row multiplications. This method employed in solving the problem has been used successfully in the problem of optimal interpolation with incomplete polynomials lacking the linear term [14]. For reasons to be explained below, however, other problems of incomplete or lacunary polynomial interpolation are more complicated and still await solution.

More general problems — the nonsingularity problem:

For some choices of the range space Y, the problem described above may or may not occur, but a second problem does arise. The matrix corresponding to (2) may indeed be reducible to the form (4), but condition (i) might not be implied by the linear independence of the functions $q_1, \ldots, q_n$ defined in (4a). This situation will occur if for example $q_1, \ldots, q_n$ are polynomials of degree greater than $n-2$.

The simplest example in which the functions $q_1, \ldots, q_n$ are of too high degree is that of interpolation into a space Y of polynomials of degree $n+m$ on an interval $[a, b]$, with each y in Y having a common factor of

$$(t - t_{n+1}) \ldots (t - t_{n+m}),$$

where $t_{n+1}, \ldots, t_{n+m}$ are points outside of $[a, b]$. For convenience, we may locate these points to the right of b, in such fashion that $b < t_{n+1} < \ldots < t_{n+m}$. This case is handled in [11].

The functions $y_0, \ldots, y_n$ used in interpolation are defined for $i=0, \ldots, n$ by

$$y_i(t) = \prod_{\substack{j \neq i \\ j=0}}^{n+m} (t - t_j)\,(t_i - t_j)^{-1},$$

and so the "cancellation" steps which reduce the matrix (2) to the form (4) are handled just as in the classical case. The functions $q_1, \ldots, q_n$ defined as in (4a) are of degree $n+m-2$ or less, which means that Proposition 1 cannot be applied. Neverthess, the functions $q_1, \ldots, q_n$ defined in this case do obey the sign properties (a), (b), and (c) described above, and it can be shown that they obey an additional sign property on certain points $T_{n+1}, \ldots, T_{n+m}$, so situated that $t_n < T_{n+1} < t_{n+1} < \ldots < t_{n+m-1} < T_{n+m}$ namely

(d) For $k=n+1, \ldots, n+m$, $q_j(T_k)\, q_1(T_k) \leq 0$ for $j=2, \ldots, n$.

The proof of this case is now completed by employment of Proposition 2:

Proposition 2: Let $t_1, \ldots, t_{n+m-2}$ and $T_1, \ldots, T_{n+m}$, be so ordered that $T_1 < t_1 < T_2 < \ldots < t_{n+m-2} < T_{n+m-1} < t_{n+m-1} < T_{n+m}$ and let polynomials $q_1, \ldots, q_n$ be given which satisfy the sign conditions (a) - (d). Then for each $k \in \{1, \ldots, n\}$, $\det(q_i(t_j))^{n-1}_{j=1}{}^{n}_{\substack{i=1\\ i\neq k}} \neq 0$.

<u>Proof:</u> We again sketch the proof of this proposition, which proceeds by induction on m, reducing to Proposition 1 if $m=0$ and to its Corollary if $m=1$. If $m>0$, it is possible to decrease m and the degree of the functions $q_1, \ldots, q_n$ without changing their values at $t_1, \ldots, t_{n-1}$ and without changing properties (a) - (d). One simply defines the polynomial $r(t)$ of degree $n+m-2$ to be zero at $t_1, \ldots, t_{n+m-2}$ and 1 at T_{n+m}, and it may be seen that for $i \in \{1, \ldots, n\}$, $q_i(t) - q_i(T_n)\, r_i(t)$ preserves the sign properties (a) - (d) on $T_1, \ldots, T_{n+m-1}$, is equal to q_i at the points $t_1, \ldots, t_{n-1}$, and is zero at T_n. Thus the degree of the polynomials $q_1, \ldots, q_n$ can be reduced by one.

In the case that Y is a space of incomplete polynomials lacking the linear term [14], the reduction of the derivative matrix (2) to the form (8) can be performed, after which an argument similar to Proposition 2 suffices to complete the proof.

<u>The present situation of optimal interpolation:</u>

The author envisions a development of more general proofs tending in the direction of combined use of component parts resembling the reduction (8) and Proposition 2. Another possibility is, of course, that an entirely new way of establishing the non-singularity property (i) may be found. In the absence of such a new discovery, however, the problems involved in generalization may be seen by considering three examples. All of them involve a departure from the real line into the complex plane, even though we presume that interpolation is occurring on an interval [a, b], where $0 < a < b$:

<u>Example 1:</u> Let Y be the space consisting of all multiples of $p(t) = t^2 + 1$ by a polynomial of degree n or less. Then the cancellation steps carried out in Proposition 2 lead to uncertain results because, while the functions $q_1, \ldots, q_n$ have zeroes on $[T_1, T_n]$ which obey conditions (a) - (c), it is not certain where the other zeroes might be.

<u>Example 2:</u> Let Y consist of all polynomials of degree $n+1$ or less lacking a term with exponent k, where $k\neq 0$, 1, or $n+1$, and $n\geq 2$. Then, just as in the solved case $k=1$, we can carry out the reduction steps which end in (8). However, Proposition 2 or its like cannot be applied. Some of the zeroes of $X_1', \ldots, X_n'$ could in the first place have been complex.

Example 3: For $n>3$, let Y be the space of polynomials of degree $n+2$ or less which lack the linear and quadratic terms.

The functions $U_1, \ldots, U_{n-1}$ are polynomials with complex zeroes, whence the entries of (8) have complex zeroes.

In view of these three examples, it would seem that the next step in extending the conjectures of Bernstein and Erdos to new range spaces might be facilitated (if possible) by some sort of extension of results like Proposition 2 to the complex plane, unless, alternatively, there appears some method of bypassing such an extension.

We conclude this communication on a positive note. Proposition 2 implies a previously unknown result about Lagrange Interpolation:

Theorem: On the interval $[a, b]$, let all of the nodes $t_0, \ldots, t_n$ remain fixed except for some consecutive set of nodes $t_{i+1}, \ldots, t_{i+m-1}$, for some applicable m. Then

$$\min_{\{t_{i+1}, \ldots, t_{i+m-1}\}} \max \{\lambda_i, \ldots, \lambda_{i+m}\}$$

occurs if $t_{i+1}, \ldots, t_{i+m-1}$ are so situated that

$$\lambda_i = \lambda_{i+1} = \ldots = \lambda_{i+m},$$

which condition occurs at a unique positioning of the nodes $t_{i+1}, \ldots, t_{i+m-1}$. Moreover,

$\min_{\{t_{i+1}, \ldots, t_{i+m-1}\}} \max \{\lambda_i, \ldots, \lambda_{i+m}\}$ is not less than the norm of optimal Lagrange interpolation with polynomials of degree $\leq m$.

References

1. S. Bernstein, Sur la limitation des valeurs d'un polynome, Isv. Akad. Nauk SSSR 7 (1931), 1025-1050.

2. C. de Boor & A. Pinkus, Proof of the conjectures of Bernstein and Erdos concerning polynomial interpolation, J. Approx. Theory 24 (1978), 289-303.

3. C. de Boor, Polynomial interpolation, in Proceedings of the International Congress of Mathematicians, Helsinki, 1978.

4. D. Braess, Nonlinear approximation theory, Springer Verlag Berlin, 1986.

5. L. Brutman, Polynomial and rational projections in the complex plane, SIAM J. Numer. Anal. 17, (1980), 366-372.

6. L. Brutman & A. Pinkus, On the Erdos conjecture concerning minimal norm interpolation on the unit circle, SIAM J. Numer. Anal. 17 (1980), 373-375.

7. P. Erdos, Some remarks on polynomials, Bull. Amer. Math. Soc. 53 (1947), 1169-1176.

8. T. Kilgore, Optimization of the norm of the Lagrange interpolation operator, Bull. Amer. Math. Soc. 83 (1977), 1069-1071.

9. T. Kilgore, A characterization of the Lagrange interpolating projection with minimal Tchebycheff norm, J. Approx. Theory, 24 (1978) 273-288.

10. T. Kilgore, A note on functions with interlacing roots, J. Approx. Theory, 43 (1985), 25-28.

11. T. Kilgore, Optimal interpolation with polynomials having fixed roots, J. Approx. Theory, 49 (1987) 378-389.

12. T. Kilgore, A note on optimal interpolation with rational functions Acta Sci. Math. (Szeged), in press.

13. T. Kilgore, Optimal interpolation in a general setting--A beginning, J. Approx. Theory, in press.

14. T. Kilgore, Best interpolation with polynomials lacking the linear term, J. Approx. Theory, in press.

15. P. Morris & E. Cheney, On the existence and characterization of minimal projections, J. Reine Angew. Math. 270 (1974), 61-76.

16. M. J. D. Powell, "Approximation Theory and Methods," Cambridge University Press, Cambridge, 1981.

Department of Algebra, Combinatorics, and Analysis
Division of Mathematics
Auburn University,
Auburn, AL 36849
USA