ON POLYNOMIAL APPROXIMATION WITH DEVIATIONS IN PRESCRIBED RATIOS

ABRAHAM SPITZBART AND NATHANIEL MACON[1]

1. **Introduction.** The standard formulas for polynomial interpolation provide a representation for the (unique) polynomial $y(x)$ of degree n such that

$$y(x_i) = y_i, \qquad i = 0, 1, \cdots, n,$$

where x_i and y_i are given, with the x_i distinct. (For a discussion of these methods, cf. [1]). The present paper considers an extended problem for which a polynomial of degree $n-1$ approximates the $n+1$ values y_i at $x=x_i$, respectively, with deviations from the given values which are in prescribed ratios. Complete results are obtained for this extended problem, with arbitrary, distinct x_i.

The results are then applied to the two important special cases (1) equally spaced points (2) a distribution of x_i determined by a Chebychev approximation.

2. **Statement of the problem.** Let x_i, y_i, λ_i, $(i=0, 1, \cdots, n)$ be given, with the x_i distinct, throughout. We seek a polynomial $y(x)$ of degree $n-1$ such that

$$(1) \qquad y(x_i) = y_i - \lambda_i d, \qquad (i = 0, 1, \cdots, n),$$

where d remains to be determined.

If we write

$$(2) \qquad y(x) = \sum_{\nu=0}^{n-1} b_\nu x^\nu,$$

then the condition (1) becomes

$$(3) \qquad \sum_{\nu=0}^{n-1} b_\nu x_i^\nu + \lambda_i d = y_i, \qquad (i = 0, 1, \cdots, n).$$

This is a system of $n+1$ linear equations in the $n+1$ quantities b_ν, $(\nu=0, 1, \cdots, n-1)$, and d. Let B denote the column vector consisting of the b_ν and d; let W be the matrix of coefficients of these quantities; and let Y denote the column vector of y's. In matrix form the system (3) becomes

Received by the editors August 15, 1957.

[1] The authors are now at the University of Wisconsin-Milwaukee and the Alabama Polytechnic Institute, respectively.

$$WB = Y.$$

Before proceeding with the solution of this system, we prove a general theorem from which a simple condition that W be nonsingular follows. A result in Polya and Szegö [2] is the special case of this theorem in which $\lambda_0 = 1$, $\lambda_i = 0$ for $i = 1, \cdots, n$.

THEOREM 1. *Let*

$$D_k = \begin{vmatrix} 1 & x_0 & \cdots & x_0^{k-1} & \lambda_0 & x_0^{k+1} & \cdots & x_0^n \\ 1 & x_1 & \cdots & x_1^{k-1} & \lambda_1 & x_1^{k+1} & \cdots & x_1^n \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 1 & x_n & \cdots & x_n^{k-1} & \lambda_n & x_n^{k+1} & \cdots & x_n^n \end{vmatrix},$$

and let Δ_{n+1} be the Vandermonde determinant obtained by replacing λ_i by x_i^k, $(i = 0, 1, \cdots, n)$, in D_k. Then we have

$$D_k = \Delta_{n+1} \sum_{i=0}^{n} \frac{\lambda_i}{p'(x_i)} \sum_{j=0}^{n-k} (-1)^j \sigma_j x_i^{n-k-j},$$

where $p(x) = \prod_{i=0}^{n} (x - x_i)$ and σ_j is the sum of all products of the x_i taken j at a time without repetitions or permutations, $(\sigma_0 \equiv 1)$.

PROOF. Let $\pi_n(x) = \sum_{\nu=0}^{n} c_\nu x^\nu$ be the polynomial of degree n such that $\pi_n(x_i) = \lambda_i$, $(i = 0, 1, \cdots, n)$. If we solve the system

$$\sum_{\nu=0}^{n} c_\nu x_i^\nu = \lambda_i, \qquad i = 0, 1, \cdots, n, \tag{4}$$

by Cramer's rule, it follows that $c_k = D_k/\Delta_{n+1}$. But c_k can also be obtained by applying Lagrange's interpolation formula to obtain $\pi_n(x)$ and then collecting the terms involving x^k, which gives for c_k the double sum in the statement of the theorem; the theorem is proved. In particular, we have $c_n = \sum_{i=0}^{n} \lambda_i/p'(x_i)$.

Setting $k = n$ we obtain the following

COROLLARY.

$$\text{Det } W = \Delta_{n+1} \sum_{i=0}^{n} \frac{\lambda_i}{p'(x_i)} = \Delta_{n+1} c_n.$$

Since the x_i are assumed to be distinct, we have $\Delta_{n+1} \neq 0$. Thus, the following theorem holds.

THEOREM 2. *The system* (3) *is nonsingular if and only if $c_n \neq 0$*, i.e., *if and only if no polynomial of degree less than n passes through the points (x_i, λ_i), $i = 0, 1, \cdots, n$.*

For later applications we state the result below.

COROLLARY. *If the* x_i, $(i=0, 1, \cdots, n)$, *are monotonic, and the* λ_i *are alternating in sign, then* W *is nonsingular.*

3. **Determination of the deviations.** We now assume that the condition of Theorem 2 is satisfied. We may then write the solution of (3) as

$$B = W^{-1}Y. \tag{5}$$

The solution of this system depends on the determination of W^{-1}, which inverse is independent of the y_i. Let w_{ij} $(i, j=1, 2, \cdots, n+1)$ denote the element in the ith row and jth column of W^{-1}. Since d, uniquely determined by (3), is the last element in the column vector B, it follows that

$$d = \sum_{j=0}^{n} w_{n+1,j+1}y_j. \tag{6}$$

Further, we can prove

THEOREM 3. *The elements in the last row of* W^{-1} *are given by*

$$w_{n+1,j+1} = \frac{1}{c_n p'(x_j)},$$

$j=0, 1, \cdots, n$. *Thus, the ratios of the elements in the last row of* W^{-1} *are completely determined by the distribution of the* x_i, *and are independent of the* λ_i.

PROOF. Rewriting (3) as $\sum_{\nu=0}^{n-1} b_\nu x_i^\nu = y_i - \lambda_i d$, $(i=0, 1, \cdots, n)$, one sees that the polynomial $y(x)$ defined by (2) takes on the $n+1$ values $y_i - \lambda_i d$ at the x_i, $(i=0, 1, \cdots, n)$. Therefore, by Lagrange's formula, it may be written

$$y(x) = \sum_{i=0}^{n} \frac{p(x)}{p'(x_i)(x - x_i)} (y_i - \lambda_i d). \tag{7}$$

Since, on the other hand, $y(x)$ is of degree $n-1$, the coefficient of x^n in this sum is zero; thus,

$$\sum_{i=0}^{n} \frac{y_i - \lambda_i d}{p'(x_i)} = 0.$$

Consequently, we have

(8) $$d = \frac{1}{c_n} \sum_{i=0}^{n} \frac{y_i}{p'(x_i)} \, .$$

Since the elements $w_{n+1,j+1}$ are independent of the y_i, the result follows immediately by comparing the above with (6).

The polynomial (7) along with the value of d given by (8) constitutes one form of the polynomial $y(x)$ sought in the original statement of the problem. Should we require the individual coefficients b_ν $(\nu=0, 1, \cdots, n-1)$, the following methods are available:

(i) We may substitute the now known value of d into any n of the equations (3) and solve the resulting system. The matrix of coefficients is a Vandermonde matrix, the inverse of which is known explicitly [**3**];

(ii) We may collect terms of the various powers of x in (7);

(iii) We may obtain an explicit representation for W^{-1} and perform the indicated matrix multiplication in (5).

We choose to develop method (iii) in detail, since the determination of W^{-1} is of interest in itself.

4. **Inversion of the matrix of coefficients.** The elements in the last row of W^{-1} are given by Theorem 3. In this paragraph, we show that the remaining elements of W^{-1} can be expressed very simply in terms of the elements in this last row and the elements of the inverse of the Vandermonde matrix $V=\{x_{i-1}^{j-1}\}$, $(i, j=1, 2, \cdots, n+1)$ of order $n+1$ [cf. **3**].

Again we assume $c_n \neq 0$, and rewrite (4) in the form

$$-\frac{1}{c_n}\left(\sum_{\nu=0}^{n-1} c_\nu x_i^{\nu} - \lambda_i \right) = x_i^{n}.$$

Thus, if we write

$$A = \begin{pmatrix} 1 & 0 \cdots 0 & -c_0/c_n \\ 0 & 1 \cdots 0 & -c_1/c_n \\ \cdot & \cdot \cdots \cdot & \cdot \\ 0 & 0 \cdots 1 & -c_{n-1}/c_n \\ 0 & 0 \cdots 0 & 1/c_n \end{pmatrix},$$

it follows that

$$WA = V;$$

and so

$$W^{-1} = AV^{-1}.$$

Let us denote the elements of V^{-1} by v_{ij} $(i, j=1, 2, \cdots, n+1)$. Performing the matrix multiplication indicated above, we get

$$w_{n+1,j} = v_{n+1,j}/c_n,$$

$$w_{ij} = v_{ij} - \frac{c_{i-1}}{c_n} v_{n+1,j} = v_{ij} - c_{i-1}w_{n+1,j},$$

for $i=1, 2, \cdots, n$; $j=1, 2, \cdots, n+1$. Finally, it follows from (4) that

$$c_{i-1} = \sum_{j=1}^{n+1} v_{ij}\lambda_{j-1},$$

for $i=1, 2, \cdots, n+1$. We shall call the quantity c_{i-1} the *λ-sum of the ith row of* V^{-1}.

The results are summarized below.

THEOREM 4. *The elements of* W^{-1} *can be obtained from those of* V^{-1} *as follows*:

(a) *the elements of the last row are given by* $w_{n+1,j} = v_{n+1,j}/c_n$, $(j=1, 2, \cdots, n+1)$;

(b) *to obtain the ith row of* W^{-1} *for* $i \leqq n$, *form the λ-sum of the ith row of* V^{-1} *and subtract the product of this λ-sum and the last row of* W^{-1} *from the ith row of* V^{-1}.

Now let k be any of the numbers $0, 1, \cdots, n$, and assume $c_k \neq 0$. Let $\{D_k\}$ denote the matrix of the D_k displayed in Theorem 1, and let d_{ij} be the element in the ith row and jth column of $\{D_k\}^{-1}$. By methods similar to those used above, one can derive the relations

$$d_{k+1,j} = v_{k+1,j}/c_k,$$

$$d_{ij} = v_{ij} - \frac{c_{i-1}}{c_k} v_{k+1,j} = v_{ij} - c_{i-1}d_{k+1,j},$$

where $i=1, 2, \cdots, k, k+2, \cdots, n+1$; $j=1, 2, \cdots, n+1$. Clearly, Theorem 4 is the special case in which $k=n$.

5. **Application to equally spaced points.** Suppose now that $x_i = x_0 + ih$, $(i=0, 1, \cdots, n)$. We show that in this case the elements of the last row of W^{-1} are simple to obtain from Theorem 3; and thus, the remaining rows come immediately from Theorem 4 and a previous result expressing the elements of V^{-1} in terms of Stirling numbers when the x_i are equally spaced [**3**].

We have, from Theorem 4,

$$w_{n+1,j+1} = 1/[c_n p'(x_j)],$$

$(j=0, 1, \cdots, n)$, where $p(x)=\prod_{i=0}^{n}(x-x_i)$, and $c_n=\sum_{i=0}^{n}\lambda_i/p'(x_i)$ is assumed to be nonzero. Since here, $x_j-x_i=(j-i)h$, one sees that

$$p'(x_j) = (-1)^{n-j}j!(n-j)!h^n = (-1)^{n-j}n!h^n \Big/ \binom{n}{j}.$$

Hence, we have

THEOREM 5. *If the points x_i $(i=0, 1, \cdots, n)$ are equally spaced, then the elements of the last row of W^{-1} are proportional to the binomial coefficients; more exactly, if $x_i=x_0+ih$, then*

$$w_{n+1,j+1} = (-1)^{n-j}\binom{n}{j} \Big/ (n!h^n c_n), \qquad (j = 0, 1, \cdots, n)$$

where $c_n=\sum_{i=0}^{n}\lambda_i/p'(x_i)$ is assumed to be nonzero.

As an illustration in which c_n may be easily evaluated, we choose the important special case for which $\lambda_i=(-1)^i$, $i=0, 1, \cdots, n$. We know from the corollary to Theorem 2 that W^{-1} exists for this case. We have

$$c_n = \sum_{i=0}^{n}(-1)^i/p'(x_i) = [(-1)^n/(n!h^n)]\sum_{i=0}^{n}\binom{n}{i} = (-1)^n 2^n/(n!h^n)$$

and, further,

$$w_{n+1,j+1} = (-1)^n\binom{n}{i} \Big/ (2h)^n.$$

For a given set of y_i we may now obtain d, as given in (6), explicitly as

$$d = [1/(2h)^2]\sum_{i=0}^{n}(-1)^i\binom{n}{i}y_i$$

6. Application to Chebychev approximation. Another important special case of the problem here considered was treated in a recent paper [4]. The points x_i, $i=0, 1, \cdots, n$ are defined by

$$x_0 = 0, \qquad x_n = 1, \qquad T_n'(x_i) = 0, \qquad i = 1, 2, \cdots, n-1$$

where $T_n(x)$ is the Chebychev polynomial of degree n for the interval $(0, 1)$,

$$T_n(x) = (-1)^n \cos[n \arccos(2x-1)]$$

and $\lambda_i=(-1)^i$, $i=0, 1, \cdots, n$. (For a general discussion of Chebychev approximation see [5, p. 197 ff]. Numerous examples are given in [6], and an application to the determination of optimum interval tables may be found in [7].) In particular, we may write x_j explicitly as

$$x_j = \sin^2 (j/2n)\pi, \qquad j = 0, 1, \cdots, n.$$

The polynomial $p(x) = \prod_{i=0}^{n} (x - x_i)$ becomes

$$p(x) = x(x-1)T_n'(x)/(n\cdot 2^{2n-1})$$

from which we obtain

$$p'(x_0) = -\frac{T_n'(0)}{n\cdot 2^{2n-1}}, \quad p'(x_n) = \frac{T_n'(1)}{n\cdot 2^{2n-1}}, \quad p'(x_i) = \frac{x_i(x_i-1)T_n'(x_i)}{n\cdot 2^{2n-1}},$$
$$i = 1, 2, \cdots, n-1.$$

From [4], we then obtain

$$p'(x_0) = n/2^{2n-2}, \quad p'(x_n) = (-1)^n n/2^{2n-2}, \quad p'(x_i) = (-1)^i n/2^{2n-1},$$
$$i = 1, 2, \cdots, n-1.$$

Here again c_n is easily computed, and is $c_n = 2^{2n-1}$. The last row of W^{-1} then becomes

$$\frac{1}{2n}, -\frac{1}{n}, \frac{1}{n}, \cdots, \frac{(-1)^{n-1}}{n}, \frac{(-1)^n}{2n}$$

and the familiar ratios $1, -2, 2, \cdots, (-1)^{n-1}2, (-1)^n$, appear.

References

1. F. B. Hildebrand, *Introduction to numerical analysis*, New York, McGraw-Hill, 1956.

2. G. Polya and G. Szegö, *Aufgaben und Lehrsätze aus der Analysis*, Vol. II, New York, Dover, 1945, Problem 10, p. 99.

3. N. Macon, and A. Spitzbart, *Inverses of Vandermonde matrices*, Amer. Math. Monthly vol. 65 (1958) pp. 95–100.

4. A. Spitzbart, and D. L. Shell, *A Chebycheff fitting criterion*, Journal of the Association for Computing Machinery vol. 5 (1958) pp. 22–31.

5. A. S. Householder, *Principles of numerical analysis*, New York, McGraw-Hill, 1953, pp. 197-200.

6. C. Hastings, Jr. *Approximations for digital computers*, Princeton University Press, 1955.

7. H. R. J. Grosch, *The use of optimum interval mathematical tables*, Proceedings Scientific Computation Forum, I.B.M. Corp., New York, 1948, pp. 23–27.

General Electric Company, Cincinnati, Ohio